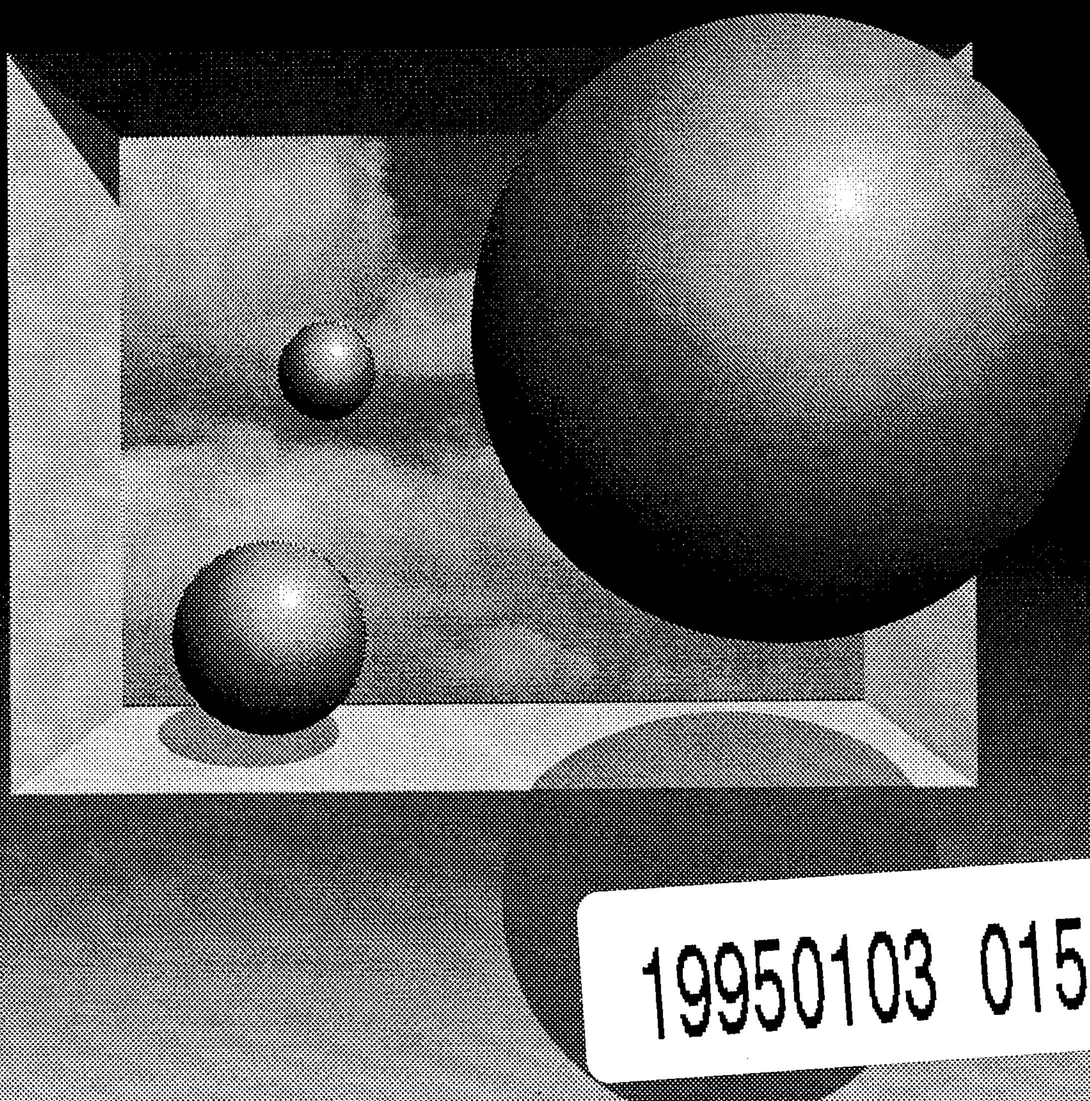
TEACHING EXCELLENCE

Methodologies to Assess Teaching Effectiveness



Maj Steven G. Webb, USAF

Research Report No. AU-ARI-93-10

Teaching Excellence
Methodologies to Assess Teaching Effectiveness

STEVEN G. WEBB, Major, USAF

ARI Command-Sponsored Research Fellow
United States Air Force Academy

Air University Press
Maxwell Air Force Base, Alabama

November 1994

Disclaimer

This publication was produced in the Department of Defense school environment in the interest of academic freedom and the advancement of national defense-related concepts. The views expressed in this publication are those of the author and do not reflect the official policy or position of the Department of Defense or the United States government.

This publication has been reviewed by security and policy review authorities and is cleared for public release.

Contents

Illustrations

Foreword

The key to success and advancement in today's Air Force is education. In light of the current force restructuring, Air Force leaders must insure that its members have the best training and the best education possible. To accomplish this quality training and education, the Air Force must develop and implement the most effective educational programs for its personnel. This feat requires examining current programs to assess their effectiveness and examining innovative programs to determine how they influence the educational process.

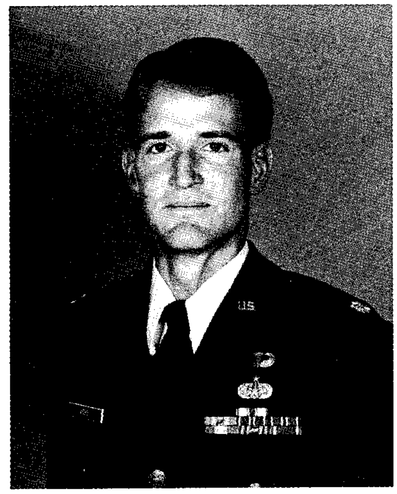
Maj Steven G. Webb conducts a comprehensive study of how to develop methodologies to assess teaching effectiveness. He presents a thorough review of many teaching styles, and then examines methodologies which can assess the effectiveness of those styles. He develops guidelines and principles to help educators develop innovative courses and integrated assessments which will provide quality learning environments for students. Finally, Major Webb applies these guidelines and principles to the development of an engineering mechanics course at the Air Force Academy.

Anyone who wishes to enhance the Air Force educational system must understand and apply the basic guidelines and principles Major Webb outlines. Doing otherwise will provide a disservice to Air Force personnel and, consequently, to the Air Force and the United States.

Robert M. Johnston

ROBERT M. JOHNSTON, Colonel, USAF
Director
Airpower Research Institute

About the Author



Maj Steven G. Webb

Maj Steven G. Webb completed this study as the United States Air Force Academy's (USAFA) command-sponsored research fellow for 1992–93 at the Airpower Research Institute (ARI), Air University College for Aerospace Doctrine, Research, and Education (AUCADRE) at Maxwell Air Force Base (AFB), Alabama. In 1980 Major Webb graduated from the USAFA with a bachelor of science degree in astronautical engineering and mathematics. He received a Guggenheim Fellowship to attend Princeton University, from where he received a master's in aerospace engineering in 1981. From 1981 to 1983, Major Webb was assigned to Los Angeles Air Force Station as an advanced space systems project engineer for the deputy of technology. In 1983 he was a member of the initial cadre that established the Air Force Space Technology Center (AFSTC) at Kirtland AFB, New Mexico. He was a program manager for advanced space technology programs at the AFSTC until 1985, when he was assigned to the Air Force Institute of Technology (AFIT) at Wright-Patterson AFB, Ohio, where he earned a doctoral degree in aeronautical engineering in 1988. Major Webb was then assigned as an instructor in the Engineering Mechanics Department at the USAFA. While there, he attained the rank of associate professor, and was in charge of the department's large space structure research program. In 1992 the Air Force Academy selected him as a research fellow at ARI and concurrently to attend Air Command and Staff College. Major Webb is presently assigned to the National Reconnaissance Office. He and his wife, Dina, have three sons, Brandon, Ryan, and Sean.

Preface

In 1992 Brig Gen Ruben A. Cubero tasked me to develop methods to assess the effectiveness of the networked classroom system (NCS), which the Air Force Academy recently installed. Indeed, this was a daunting assignment. Although I taught a variety of engineering courses at the Air Force Academy, I never devoted much time to exploring innovative teaching styles or to investigating diverse methodologies to assess student performance in a learning environment. While researching this report I gained a tremendous respect for educators who sincerely wished to enhance the undergraduate educational process.

I believe teaching is one of the most difficult and frustrating yet rewarding professions imaginable. I am convinced that educators must continually search for ways to enhance the learning environment for students. An instructor simply cannot decide to teach students a topic; he or she must devote much time and energy to develop a course that not only addresses the instructor's efficacy but also enhances the students and their learning habits and attitudes. This is a challenge, and educators have many conflicting views on how to approach teaching. It is imperative, though, that the instructor considers the students first when choosing a particular teaching style or learning environment for the course. Coincidentally, the instructor must integrate assessments into the course; an instructor cannot add them as an afterthought once he or she has developed the course. Assessments permit the instructor to insure that goals and objectives for a course are realistic, achievable, and measurable.

This study is not all encompassing; in fact, I do not mention a number of teaching styles, innovative educational technologies, and assessment methodologies. I never intended this study as a treatise on education; rather, I hope the reader uses the information and the guidelines and principles I offered in chapters 5 and 6 to develop effective courses while considering the learning process of the students.

I wish to express my deep gratitude to Col Cary Fisher, Lt Col Bob Pieri, and Lt Col Ray Richardson for helping me to get this research fellowship; it is refreshing to know Air Force leaders who look after their people. I am indebted to Lt Col John Souders, Dr Kathleen Hannafin, and Dr Michael Hannafin for the comments, suggestions, and expertise they provided. Their enthusiasm towards education was contagious, and they contributed tremendously to my pursuit of a quality project. I also want to thank my research advisor, Dr Karl Magyar, and my editor, Dr Richard Bailey, for their valuable assistance in completing this document.

My deepest appreciation is reserved for my family. This year has been a difficult, yet rewarding, one: attending Air Command and Staff College while working on this study has made for a long year with little time for my family. I cannot say enough good things about my wife, Dina. She cheerfully picked up my family duties when I had to study or write and provided much needed inspiration and encouragement. I never thanked her enough for all her help and support or never adequately expressed my appreciation to her, but I could not have completed this document without her. Finally, I thank my sons, Brandon, Ryan,

and Sean, for their patience and understanding this past year; I hope that they understand why I could not spend much time with them. If people find this document worthwhile, then the sacrifices my family made for me will have been worth it.

Introduction

The only thing constant about change is that it is inevitable. Nothing could be more true about this statement than how it so aptly defines society today. The data base of knowledge acquired by the human race is accelerating at an astonishing, yet sometimes exhilarating, pace. The technological marvels being introduced today were mere dreams a few years ago. Technology is a large part of our very existence; it is responsible for our standard of living. In fact, technology plays a role in every facet of our lives.

The field of education is one area, however, that technology has not completely permeated; that is, until recently. Ever since people undertook the task of educating others, they have imparted knowledge through traditional means: lectures, discussions, demonstrations, and tutoring. Recently, though, technological advances have made it possible to use such innovative forms of media as interactive videodiscs and computer systems to teach people in a fashion different from the traditional methods. Education in the Air Force has paralleled this evolution of teaching. Considering the importance of education in the Air Force, one is not surprised that the Air Force devotes a great deal of time, energy, and resources to improving the education of its personnel. In fact, the Air Force Academy recently developed a networked classroom system (NCS) to explore the potentials of innovative educational technologies.

However, who can say that using innovative educational technologies to teach students is better than the traditional methods of teaching? Before one determines which style of teaching is better, one should compare the performance of students enrolled in courses using innovative educational technologies to the performance of students enrolled in similar courses taught in the traditional manner. Thus, educators should employ a methodology, or several methodologies, to assess student performance to facilitate comparisons and to examine the advantages and disadvantages of both types of teaching styles.

Similarly, educators have advanced novel styles of teaching which do not rely on innovative educational technologies. These styles also may enhance a student's learning process, and educators must determine if they are more effective than the more traditional teaching methods or, for that matter, than the innovative educational technologies. Again, the NCS provides a controlled environment to investigate the effectiveness of a plethora of teaching styles.

In fact, the Air Force Academy plans to compare different styles of teaching within the context of the NCS and to investigate how those teaching styles enhance the education of its cadets. Such studies of various teaching styles, including the use of innovative educational technologies, on student performance in the classroom will enable the Air Force Academy and other Air Force, government, and civilian educational institutions to improve their learning environments.

Educators can identify three stages of assessing teaching effectiveness. First, they need to examine existing methodologies researchers use to assess student performance. One or more of these methodologies may prove to be useful in comparing various styles of teaching. If not, the educators can use the second stage of this process to develop new methodologies to assess the effectiveness of those

teaching styles. Third, they must implement the chosen methodology, or methodologies, during actual courses to assess the effectiveness of the different teaching styles instructors use to enhance the students' learning process.

This research report addresses only the first stage of the process outlined above. I first establish parameters which frame a controlled and an unbiased assessment of both the traditional methods of teaching and teaching using innovative educational technologies. I then examine existing methodologies researchers use to assess the effectiveness of different teaching styles. Finally, I recommend guidelines and principles to follow when educators develop comprehensive methodologies to assess student performance and the effectiveness of different teaching styles accurately and thoroughly.

However, before I discuss any teaching style or assessment methodology, I must first define the scope of the problem. Chapter 1 does this by addressing the constraints and bounds relevant to the task of assessing teaching styles. Chapter 2 briefly reviews the traditional methods of teaching and discusses various aspects of innovative educational technologies. Chapter 3 addresses existing comparisons of traditional teaching and teaching using educational technology. Chapter 4 identifies methodologies which researchers have used to assess student performance and the effectiveness of various teaching styles. Chapter 5 recommends guidelines to follow in conducting assessments of teaching styles, and chapter 6 outlines basic principles which educators must address to develop assessment methodologies. Finally, chapter 7 provides a case study for the Air Force Academy that applies the guidelines and principles discussed in chapters 5 and 6 to the development of an undergraduate engineering course and its associated assessment methodologies.

Chapter 1

Constraining the Study

As I stated in the introduction, this research project establishes guidelines and principles for developing methodologies evaluators can use to assess the effectiveness of different teaching styles. At first glance, the scope of this research project appears rather limited and simplistic. In reality, however, this topic is quite complex, comparable to walking through a minefield without knowing where the mines are located. Emotions seem to run high when educators discuss the possibility of using novel teaching styles or innovative educational technologies to teach students. It appears that everyone has an opinion of the best way to teach and wants to defend that position. Similarly, methodologies to assess styles of teaching are abundant and have legions of supporters ready to show how superior their preferred methodology is to others.

In this report I provide an unbiased view of these fields of teaching and assessment. It is not my position to determine the "best style of teaching" or the "best methodology." Rather, I review the state of professional expertise on this subject and suggest the essential requirements necessary for thorough and unbiased assessments of teaching styles.

Some Definitions

Before proceeding further, I must define some terms which appear throughout this report. First, I must define teaching: *Webster's II New Riverside University Dictionary* defines *teaching* as imparting knowledge or skills, causing to learn by example or experience or by advocating or preaching.[1] Educators teach students in a variety of ways; in this report I refer to a particular way of teaching as a "teaching style." This style provides a logical approach to teaching students and includes a chosen strategy, the required materials, and the equipment educators use to teach. I divide teaching styles into two basic categories: traditional teaching methods that use innovative educational technologies as teaching aids. These categories are mutually exclusive because traditional teaching methods do not use innovative educational technologies and, in many cases, any supporting equipment. Educators choose, from one of these two categories, a teaching style that is appropriate for the instructor, the students, and the course; quite often, the best approach to teaching blends teaching styles.

Essentially, teaching revolves around answering three basic questions: What is an instructor to teach? How can the instructor best teach? and How well did the instructor teach? To answer the first question, the instructor must establish goals and objectives for his or her course. Answering the second question requires that he or she must develop an appropriate style of teaching for the course. Answering the final question necessitates that the instructor evaluate his or her performance and the effectiveness of the teaching style. Teaching, then, is a continuous and iterative process; instructors must constantly evaluate their learning environments and incorporate appropriate changes to their goals, objectives, and teaching styles to insure that they are providing the best possible environment for their students.[2]

The style of teaching embraces just one aspect of a learning environment that instructors expose students to during their learning process. This learning environment also possesses such physical attributes of the classroom as the environmental conditions, desks, chairs, classroom arrangement, and equipment that is available for the instructor's use. The materials the instructors and students use comprise part of the learning environment. Finally, the instructors' and the students' interests, attitudes, motivations, and aptitudes influence the learning environment.

Within this environment, educators either instruct or teach students. While instructors sometimes use instruction and teaching methods interchangeably, the methods do differ; students process information externally during instruction while they process information both internally and externally during teaching. In other words, the goal of instruction is to disseminate basic knowledge and skills so that the students can accomplish specific tasks. Conversely, teaching combines instruction with developing a student's critical thinking and insight. For example, students can receive instruction on how to assemble a particular product, because during the actual teaching process students not only learn how to produce the product, they also develop insight into how and why it works.[3] When I refer to learning in this report, I include teaching, not solely instruction.

Teaching insures that students achieve the desired learning objectives of the Course. Educators usually identify two types of objectives for the course: behavioral and cognitive. Behavioral objectives normally result from instruction; they require students to master discrete information through such means as listing and identifying content-bound facts. Students achieve these objectives by overtly demonstrating knowledge and skills. Conversely, teaching helps student to achieve cognitive objectives. They develop an understanding of the material presented during class by retrieving prior knowledge and integrating and synthesizing information to gain new insight into the information and to relate it to other concepts. Cognitive objectives entice the students to develop higher order learning skills, abilities, and understanding.[4]

To measure these behavioral and cognitive learning objectives, educators must develop assessment methodologies. These methodologies include

procedures and methods to collect and examine data. Educators then evaluate this data; evaluations require that educators use insight and sound reasoning to determine the significance of the data and to judge the results. According to Andrew Beale, "Evaluation [can be] viewed as a means of acquiring and processing the evidence needed to determine the student's level of learning and the effectiveness of teaching."[5] Educators must therefore assess such information as test results, reports, and laboratory exercises so they can evaluate the effectiveness of a particular teaching style.[6]

Similarly, students and instructors can evaluate the effectiveness of an instructor or a style of teaching by completing evaluation forms, surveys, and questionnaires. These evaluations require that the raters judge specific attributes of the instructor or teaching style; therefore, the rater conducts an evaluation rather than an assessment. Educators can use the information from these evaluations, along with their assessments of student performances, to judge the effectiveness of a specific teaching style.

In this report I refer to assessing student performances in learning environments and assessing styles of teaching. The assessments I examine and the guidelines and principles I develop address the effectiveness of different teaching styles, including both traditional teaching methods and teaching using innovative educational technologies. Although the style of teaching is crucial, educators must assess the other elements in the learning environment to insure that they effectively use the teaching style they have chosen. This report addresses many of those elements by examining how they relate to different styles of teaching; the guidelines and principles I outline incorporate these elements as essential components of thorough assessments of teaching effectiveness.

The Networked Classroom System

The Directorate of Education at the United States Air Force Academy is one organization which can use the guidelines and principles I develop to aid their studies of the effectiveness of different teaching styles. The Air Force Academy recently designed and built a networked classroom system (NCS) which they will use as a "pedagogical incubator."[7] The directorate wants the NCS to serve as a testbed for teaching and learning styles, to examine innovative educational technologies, and to act as a laboratory for faculty development.[8]

The Air Force Academy began using the NCS during the fall semester of 1992. International Business Machines (IBM) installed 18 networked microcomputer workstations, each with six megabytes (MB) of random access memory and a 120 MB disc drive. In addition, IBM installed two "educator" workstations which control such peripheral equipment as a videocassette recorder, a videodisc player, an overhead projector, a projection television, and a video camera. Instructors use IBM's Multimedia Educator/2 System (MME)

to present course materials manually or automatically to the students on the network. This system also allows them to develop and edit their own course software to use with the NCS. Finally, instructors can use these workstations to control the software and data students load at their student workstations. The capabilities of the NCS and the MME permit educators to develop effective styles of teaching for their courses.[9]

If it fully utilizes the NCS, the Air Force Academy can, because of its unique nature as a specialized undergraduate educational institution, produce a watershed of studies on teaching styles and their effectiveness in enhancing and increasing student learning. One attribute of the Air Force Academy that distinguishes it from other postsecondary institutions is its capability to establish distinct control and experimental groups with which educators can accomplish studies of teaching. In-depth data on the various abilities of cadets is readily available, and educators can easily manipulate cadet class schedules to insure that they have distinct groups of students on which to conduct their studies. When compared to the activities of other postsecondary institutions, the extracurricular activities in which all cadets participate are fairly uniform at the Air Force Academy, and the effect of these activities on each cadet's time in class is minimal. Further, it is a simple matter to trace the cadets in both the control and experimental groups in future semesters and to conduct follow-up assessments.

A second attribute in studying and assessing the effectiveness of different teaching styles is that the primary duty of Air Force Academy instructors is to teach the cadets. Unlike many other undergraduate institutions, the instructors do not have to devote a large part of their time and energy to research projects. Consequently, with few exceptions, instructors consider teaching as their full-time profession. This consideration makes them amenable to experimenting with various teaching styles, and they are constantly investigating new ways to teach cadets.[10]

Therefore, after looking at the tremendous capabilities and potential for research the NCS and the Air Force Academy possess, educators may be tempted to jump in and begin conducting studies to assess the effectiveness of different teaching styles. However, it is extremely crucial to step back at this point and define the boundaries of such studies. Otherwise, this report and those that follow will attempt to generalize results based on arbitrary baseline criteria, and, as a result, make suspect any studies researchers generate. I must therefore first define or specify a number of variables and constraints to insure that educators responsibly undertake any study to assess the effectiveness of teaching styles. Hence, the remainder of this chapter identifies those variables and constraints that establish the baseline criteria from which educators can assess styles of teaching.

I must emphasize that I will not address an assessment of the NCS itself. Dan Surry, a doctoral student at the University of Georgia, has developed some techniques to evaluate various technical and physical aspects of this classroom. They include the classroom's comfort level, opinions of students and instructors relating to the classroom's physical attributes, and

advantages or disadvantages of the classroom's capabilities.[11] Instead, I will discuss how educators should assess the use of teaching styles, including using traditional teaching methods and teaching using innovative educational technologies, to increase the effectiveness of teaching.

What is Teaching Effectiveness?

The first, and probably the most important, task for educators is to define *teaching effectiveness*. Richard Shavelson, Noreen Webb, and Leigh Burstein define teaching effectiveness as the ability of the teacher to produce higher-than-predicted gains on standardized achievement tests.[12] In other words, a style of teaching is effective if the students, averaged over the whole class, answer more questions correctly on standardized tests than expected, based on pretest performances. By this standard, the students learn more than anticipated, and, therefore, the teaching style is effective.

Unfortunately, this definition of teaching effectiveness is based on several potentially incorrect assumptions. First, the standard assumes commonality of curriculum goals, objectives, and content coverage across all classes; there is no leeway for individually tailored classes. Second, the results are strictly summary in measuring the students' knowledge of the subject matter. It's not what the students know or do not know that matters, but what is important is the accumulated quantity of their knowledge in comparison to a pretest standard. The result of the student's performance is equated with knowledge in the subject area and does not account for partial knowledge, test-taking skills, or guessing abilities. Finally, this definition of teaching effectiveness is strictly aggregative across students within a class, regardless of the distribution of performance. The class average is the class performance, even though this definition fails to consider internal class variability in student performance. Educators miss this important aspect of student learning when they use summary scores of standardized tests to measure teaching effectiveness. Teaching affects both the structure and the cognitive content of student knowledge, and standardized tests cannot adequately reflect changes in a student's cognition.[13]

Clearly, educators need another definition of teaching effectiveness. Perhaps they would find it helpful to ask a fundamental question: What is meant by effectiveness? Webster defines *effectiveness* as "producing or designed to produce a desired effect."[14] In the context of education, the efficacy of teaching is to increase a student's performance in a course. Thus, a style of teaching is effective when it increases a student's performance.

The first issue, however, seeks to define what is meant by student performance. Typically, in a classroom environment, educators can directly relate student performance to the results of tests and other forms of assessments. Educators must specify how they plan to measure increases in student performance. The obvious measures are the universal performances

of students based on established criteria, such as increasing a student's level of knowledge by 10 percent or having a student achieve a minimum of 90 percent correct on standardized tests. However, this assessment offers too narrow a view of student performance. Other facets of performance concern student interest and involvement in the course; educators can subjectively measure these facets with surveys, questionnaires, and interviews. Hence, at the outset of any assessment, educators must clearly define what they consider student performance to be and how they plan to measure it.

Educators can determine a student's performance in a course in a logical sequence. Intuitively, they should begin first with a baseline scenario and state what is known prior to the course's start. The educators subsequently track the course and the student throughout the course. At the course's conclusion, they measure the performance of the student to determine the difference from the baseline scenario. This difference, then, measures the effectiveness of the style of teaching.

Unfortunately, several other questions are now apparent. First, what baseline scenario should educators use for such a study? Additionally, what criteria will they assess to determine a student's performance and, hence, the effectiveness of a particular teaching style? The latter part of this chapter addresses the second question; answering the first question necessitates defining a baseline scenario which considers the course, the students, and the instructors.

Parameters Bounding Studies of Teaching Styles

Any valid comparison of different teaching styles requires that instructors teach the same course with each style of teaching the educators wish to examine. Given this baseline criterion, researchers already have conducted a number of studies to compare the performance of students enrolled in identical courses. In fact, in several studies instructors have used innovative educational technologies to teach students, while other instructors have employed traditional teaching methods. However, Richard Clark and Gavriel Salomon reported that although cases exit where the introduction of innovative educational technologies dramatically changed student achievement or ability, "It was not the medium per se that caused the change but rather the curricular reform that its introduction enabled."[15]Educators expect this outcome, as the use of innovative educational technologies creates a different approach to teaching; the teaching model then changes. Instructors must therefore accept the fact that they will be changing their teaching model when they introduce innovative educational technologies in the classroom.

Still, instructors can teach the same course by using different pedagogies; what educators require is concrete knowledge of these desired teaching models before they develop the course.[16] For these educators, therefore, the

first step in developing a course is to identify clear, achievable, and observable cognitive goals which encompass the purpose of the course. These goals must relate to the desired levels of learning, possibly in terms of Bloom's taxonomy of the cognitive domain, which identifies six levels of learning: knowledge, comprehension, application, analysis, synthesis, and evaluation.[17] Educators may use other models of learning; the emphasis here is on insuring that the cognitive goals, rather than the technology, drive the course development, and, therefore, the pedagogical model.

Implicit in specifying the learning level for each of the cognitive goals is the need to identify the associated type of learning, whether it is short-term performance in class or long-term transfer of knowledge, skills, abilities, and understanding to related courses. Both types of learning require tight constraints on the baseline criteria; the former examines the performance of students in only one class and could be perceived as being a snapshot in time with respect to the effectiveness of the teaching style. The second type of learning necessitates using the first course in a series of related courses and following the students through the sequence. As was previously stated, educators can accomplish this at the Air Force Academy, thereby permitting studies of the long-term effects different styles of teaching have on student performance.

Once the educators have established the direction for the course, they can delineate specific, realistic, and measurable behavioral and cognitive objectives.[18] The behavioral objectives must specify concrete outcomes of the course and definitive levels of student competency, which is a prerequisite to developing appropriate teaching models to use for the different styles of teaching.[19] Educators must also specify cognitive objectives if the goals include higher order learning outcomes.[20] They must develop a single set of course behavioral and cognitive objectives for all of the styles of teaching they wish to assess, and they must specify the desired objectives in terms of only those behaviors and cognitions which they can strictly define, observe, and measure. As they develop these objectives, educators must, according to John Carroll, "Seek mainly to achieve equality of opportunity for all students, not necessarily equality of attainment."[21] In other words, educators must provide appropriate, but not always equal, opportunities for each student to learn by determining objectives which push that student's potential to his or her upper limits. The objectives should recognize that students possess different levels and different kinds of aptitudes and abilities, and the corresponding educational programs should allow the students to realize their capabilities for learning.[22]

After instructors define the behavioral and cognitive objectives, they can use them to measure student performance. However, the question remains as to how educators will measure the objectives. They must either find tests or develop tests to measure these objectives, and they must judge how appropriately the tests measure a student's ability to perform at the desired learning levels.[23] Before educators can actually begin to assess the student's performance, they must first establish impartial, thorough, and accurate means of measuring that performance.

Thus far I have specified developing teaching models appropriate to the styles of teaching instructors wish to use. Based on these models, the instructors develop a course using clear behavioral and cognitive objectives with specific desired levels of learning. Educators measure student performance against those learning levels. This approach appears to serve as an adequate baseline to compare different teaching styles. But the course itself is only one of many variables to consider. The students enrolled in the course introduce a variety of variables which educators must identify and constrain to the fullest extent possible.

Identifying Students for the Study

To compare students in a course where instructors use various teaching styles, educators must divide the students into equal groups; instructors then teach each group by using a different style of teaching. While the following discussion applies to any teaching style, the remainder of this chapter postulates that educators wish to compare the effectiveness of a traditional teaching method to the effectiveness of teaching using a particular innovative educational technology. Thus, instructors will teach students in the first, or control, group using the traditional teaching method, and they teach students in the second, or experimental, group with the aid of an innovative educational technology. If the baseline characteristics of both groups are identical, the differences between the two groups at the end of the course should be due only to the influence of two factors: the impact of the innovative educational technology on the experimental group, and the instructors for both groups. Hence, in choosing the students for the two groups, educators should base their decisions primarily on equalizing, between groups, the characteristics of those students. However, educators must realize that each student's performance in class will differ; as such, they cannot make conclusions based on the results of the group's performance taken as a whole without considering individual performances.

Volunteer effects play a critical role when educators choose students to participate in a study comparing different styles of teaching. If the educators require students to give consent to participate in the experiment, they introduce a risk of experiment reactivity: some students who know they are in an experiment tend to change their behavior.[24] One way to determine if this knowledge indeed creates a problem for the study is to randomly divide the students who volunteer into two subgroups. One group participates in the study, while the other is told they are not needed. Form a third subgroup of students who are not volunteers. Then divide each of the three subgroups of students into the experimental and control groups. Compare the performance of the subgroups with the nonvoluntary students being a baseline to measure the volunteers. If, within both the experimental and control groups, the performance of the students in the subgroups differs substantially, then there exists a volunteer effect.[25]

Forming these subgroups requires in-depth knowledge of each student taking the course the educators wish to examine. The ideal students to use in assessing different educational environments would be identical in every aspect. However, human nature dictates that every student is different, and, therefore, the groups of students enrolled in courses employing different styles of teaching will differ. To fairly compare the various teaching styles, educators must describe the students by specifying as many of their characteristics as possible. A multitude of characteristics of the students affects their learning process, and, therefore, influences their performance in class. John Biggs lists the more important of these as:

1. general abilities, such as those which educators can measure with intelligence quotient tests
2. special abilities and competencies
3. prior knowledge of the subject
4. interest in the particular subject matter
5. age
6. experience
7. general conception of learning
8. approach to learning.[26]

Since educators must identify appropriate student characteristics so they can assemble identical control and experimental groups, they must quantify the above characteristics, where possible, for each student within the groups. The educators can quantify most of these characteristics through pretests, interviews, surveys, and questionnaires; they can use the resulting information to divide the students into two nearly identical baseline groups.

Biggs also identifies other factors which affect student learning abilities:

1. structure of the course; for example, if it is either a core or an elective course for the student
2. class load of the student
3. available time for learning
4. student perception of the classroom environment
5. outside sources of stress.[27]

These student characteristics are more difficult to quantify. However, because of their ability to monitor student course loads at the Air Force Academy, educators can easily quantify and account for the first three characteristics when identifying students for both groups. Further, with a few exceptions, the classrooms at the Air Force Academy are nearly identical, which minimizes external effects on student learning due to the classroom environment. If they use the NCS, educators can employ the techniques Dan Surry developed to quantify the differences between the environment in that class and those of regular classrooms.[28] The final factor, that of outside sources of stress, is nearly impossible to quantify and therefore must remain as one uncontrollable variable.

Closely coupled with the attempt to quantify outside sources of stress is the need to identify student biases that might influence the outcome of a study that compares the effectiveness of different teaching styles. For example, an instructor ideally wants the students to put forth maximum effort in his or her class; unfortunately, unmotivated students may reduce the overall class performance if they exhibit forms of behavior that negatively influence the results of the assessments educators conduct on those individuals.[29] One way to identify these unmotivated students prior to the start of the experiment is to screen them through pretests, interviews, questionnaires, and surveys.

Another bias to consider is the student's desire to learn. Noel Entwistle notes that a student's understanding of material only partly depends on direct teaching, as students spend a great deal of time in independent study.[30] While educators find it relatively easy to conduct surveys to quantify student study time outside class, they realize how difficult it is to specify the quality of that time. When studying, students tend to have different goals, intentions, and abilities which relate to four educational outcomes: academic, vocational, personal, and social. The first three outcomes are "directed intrinsically towards aspects of the course content or extrinsically toward the qualifications or satisfactions incidentally associated with that content."[31] The fourth outcome, social, is usually not related to the course in any respect. Each student forms an informal and internal study contract so he or she knows what goals to aim for, and the student can qualify his or her relative satisfaction in pursuing those goals. Thus, the student's effort in independent study relates more towards his or her progress in fulfilling his or her own contract than it does in satisfying formal course requirements.[32] Of course, this abstraction is nearly impossible to quantify and is an uncontrollable bias when instructors identify students for the experiment.

Still another bias which is difficult to quantify deals with how students learn. Entwistle observes that a student's predominant educational orientations appear to strongly influence both the effort put into studying and its quality. Further, students matched with learning materials complementing their own style of learning progressed at a faster rate and did so more fully than students who were mismatched.[33] Coupled with the students' educational orientations are their impressions of using innovative educational technologies as an aid in their learning process. Some students are comfortable having media, like computers, videodiscs, and movies, as part of the learning environment; others prefer to learn by traditional methods and believe that they learn better without the help of innovative educational technologies. Further, students possess a wide range of comfort levels and expertise in using computers. This range may help or hinder the ability of a student to learn when instructors use computers as a teaching aid.[34]

Educators can quantify some of the above biases, such as experience in using computers, by tailoring surveys, interviews, questionnaires, and pretests to determine specific aspects of a student's ability and desire to learn. Obviously, many other biases and characteristics help to make each student unique. If the goal of forming control and experimental groups is to make the

entering students as equal as possible at the beginning of the course, educators must establish an extensive selection process. In implementing a study comparing styles of teaching, the educators must define those student characteristics which they can quantify; this step is the first one in selecting students to participate. These characteristics must be specific and concrete, and they must relate to aspects which influence a student's learning process. Surveys, questionnaires, interviews, and pretests are probably the best tools to use to quantify those characteristics, and they permit educators to form an accurate picture of each student.

However, educators must accept the fact that the picture will be incomplete; they cannot quantify many student characteristics, at least not reliably. Hence, an important step in the student selection process is to identify those characteristics which educators cannot quantify. Any study that uses control and experimental groups must clearly state, at the outset of the study, which baseline characteristics educators can define and which ones they cannot; this decision will limit the extent to which the educators can make conclusions from the study. Further, in a general sense, educators must track these undefinable characteristics throughout the study. If the student performance, as a whole, is differentially and systematically influenced by a factor other than what is intended, the results of the experiment are biased. Three methods can identify and possibly reduce or eliminate this bias. A panel of experts can judge the results to determine their freedom from bias. Educators can conduct statistical studies to measure the bias and develop correcting factors. Finally, educators may have to limit their interpretations of the results to only those justified by strong and valid evidence.[35]

Once educators identify each student's characteristics as fully as possible, they can divide the students into the control and experimental groups. The educators must accomplish this division so that the "composite" student in the control group is nearly identical to the "composite" student in the experimental group. Composite means that the quantified characteristics of each student within each group are averaged to form a composite student. While the educators will examine the study's results based on the students' performances as a group, they also must track individual performances. Within both groups some students may have significantly different performances when compared to the groups' norms. Educators should closely examine these students to identify contributing factors to their performances and determine if their performances, when considered individually, would change the study's results that the educators based on the groups' performances.

Contributing Factors Due to the Instructors

Educators also must accomplish a similar analysis of the instructors for the course they use for the study. As they did with the students, the educators must identify instructor characteristics and specify corresponding limitations, based on those characteristics, to a thorough study of the effectiveness of different teaching styles. The identified characteristics, as well as the

known limitations of the instructor, form a set of constraints around which educators can conduct this study. Since an instructor's influence on student performance can be tremendous, the educators therefore must quantify this influence where possible. However, the idea of teacher influence is quite vague, and in many situations it may be difficult to quantify.

For example, the interactions between instructors and students are quite complex. Student attitudes towards the instructor, the course, and the learning process influence the instructor's perceptions towards the students. These perceptions may then bias the instructor's attitude towards teaching the students. If the instructor notices that a student demonstrates little desire to learn the material, the instructor may either ignore or pressure the student. The resulting interactions will affect the student's performance.[36]

An instructor's experience and confidence in teaching also influence student performance. Exposing an instructor to two different teaching styles and requiring that instructor to teach a single course two different ways will probably bias the students' performances. The instructor who is more experienced using a traditional teaching method than using an innovative educational technology as a teaching aid might consequently feel more comfortable with the traditional teaching method. This phenomenon may have two opposing effects: either the students being taught in the traditional manner may receive a better education, or the instructor may place more emphasis on teaching using innovative educational technologies to the detriment of the students he or she teaches in the traditional manner.

It is quite possible, though, that using an innovative educational technology in a classroom environment may act as an equalizer for the instructor. Educational technology can make the actual teaching of students more focused and more streamlined for the instructor, since by using the technology as a teaching aid the instructor doesn't necessarily directly impart knowledge to the students. Rather, the instructor acts as a coach, facilitator, or mentor to guide students through the process of deriving appropriate conclusions from an activity.[37]

Using innovative educational technologies in teaching, however, requires that the instructor devote more time and energy towards creating coordinated and integrated classroom activities and necessitates more pedagogical awareness and expertise than do the traditional methods of teaching.[38] Furthermore, according to Michael Hannafin, innovative educational technologies can enhance the skills of the typical instructor and provide for teaching parity by allowing for the same quality of teaching in the classes that use the technologies.[39] Consequently, an inexperienced, or even a mediocre, instructor using innovative educational technologies may provide the same quality teaching imparted by an instructor who is quite effective in teaching with traditional methods. Of course, the same inexperienced instructor probably would not be the best person to teach in the traditional manner either.

Similarly, innovative educational technologies may make the good instructor even better by allowing the instructor to use teaching styles which incorporate technologies that enhance his or her teaching. Innovative educational technologies also may liberate the instructor from time-intensive tasks which may not benefit students equally; instead, the instructor could use the technologies as teaching aids, permitting him or her to devote more time to individual students.[40]

Closely tied to instructor experience and confidence level in teaching is that of the instructor's attitude towards teaching when using either the traditional methods or innovative educational technologies. For example, a possible scenario concerns an instructor who totally opposes using educational technology and creates a negative atmosphere in the classroom where educators use innovative educational technologies. Instructors with bad, or good, attitudes towards either learning environment will probably convey those attitudes in the course of teaching the students. Students exposed to those attitudes will probably reflect them in their performances in the class and on tests.

The above points demonstrate the difficulty of defining instructor quality. However, for the same reason that educators feel obligated to characterize the students before conducting a study on the effectiveness of different teaching styles, they feel an equal compulsion to characterize the instructors teaching the students. Identifying and quantifying as many instructor characteristics as possible constrain the number of unknown variables in the study and place limitations on what the study will determine.

Before evaluators can specify these characteristics, they must identify the instructors who will participate in this study. One possible way to conduct a study to assess the effectiveness of different teaching styles is to keep the instructors ignorant of their involvement. While this strategy may be possible at other undergraduate institutions, it is impossible at the Air Force Academy. Most, if not all, courses are taught in an environment where each instructor teaching a course is familiar with what the other instructors of the course are doing. Instructors would note any significant variance in the teaching style or learning environment, and they would invariably suspect that researchers were conducting an experiment without their knowledge. Consequently, educators should inform the instructors of the study; but by doing so, the educators must contemplate requesting volunteer instructors to participate in the study. Instructors who do not volunteer may introduce bias in their teaching by expressing negative attitudes towards their participation in the study, while volunteers may try their best to promote positive attitudes.

Therefore, instructors should volunteer to participate in studies assessing styles of teaching. Unfortunately, these instructors will possess a wide range of teaching experiences and confidence levels. Insuring that the studies measure the effectiveness of the teaching styles and not the effectiveness of the instructors themselves requires equivalent levels of learning for the different teaching styles educators are studying. One issue, which chapter 5 addresses, is whether a single instructor should teach two different classes,

employing traditional methods of teaching in one and innovative educational technologies in the other, or if one instructor should teach in the traditional manner while another instructor uses an innovative educational technology as a teaching aid. Both options require knowledge of an instructor's level of experience and confidence in addition to the instructor's attitudes.

It is therefore imperative to identify each instructor's characteristics related to teaching before the instructors are selected to participate in the study. Educators must specify the characteristics central to defining how experienced and confident an instructor is when teaching, either by the traditional methods or when using innovative educational technologies. Even though educators can assume volunteer instructors have positive attitudes, they must still identify instructor attitudes towards teaching in either environment. They can use surveys, questionnaires, interviews, and classroom observations to quantify each instructor's experience, confidence, and attitudes.

Educators also must determine how the instructor influences student performance in the learning environment. One method that tracks instructor influence on student performance requires that each instructor maintains a logbook or a diary during the course. The instructor's experiences and impressions of each lesson and of the students' attitudes and performances can provide insight into his or her experience and confidence levels, in addition to his or her underlying attitudes towards the course and students. The instructor's colleagues, when trained as evaluators, also can observe interactions between the students and instructor during the class and gauge the instructor's influence on the students. Finally, educators can use students' surveys and questionnaires to judge their perceptions of the instructor and his or her influence on the students.

How Do You Assess Studies of Teaching Styles?

Up to this point, I have established a baseline scenario for studying the effectiveness of various teaching styles by concentrating on three physical attributes of teaching: specifying a single set of course goals and objectives at the onset of course development; identifying and quantifying the desired baseline student and teacher characteristics; and identifying the limits of this study based on those student and teacher characteristics which educators cannot specify or quantify. Identifying, characterizing, and constraining the course, students, and instructors allow educators to establish a baseline to determine the effectiveness of different teaching styles. Let us now proceed with the discussion of the criteria to use in assessing a student's performance, which in turn will allow educators to assess the effectiveness of different teaching styles.

In discussing teaching effectiveness, I believe it is appropriate to define what assessment means. One can view assessment as the systematic

collection and analysis of descriptive and judgmental data necessary to provide useful information for decision making. Assessment serves as an information-gathering technique to appraise the quality or performance of the item educators are assessing.[41] Michael Patton states there is no one best way to conduct an assessment.[42] Still, educators must develop a rational method of assessing the effectiveness of different teaching styles. Before such a logical assessment methodology can be formulated, educators must answer six basic questions:

1. Who is the assessment for?

2. What should be the focus of the assessment?

3. What measures of objectives are possible to collect, and which are worth collecting?

4. How will educators use the assessment data to make judgments and decisions?

5. What will educators know after the assessment that they don't already know?

6. What will educators do after the assessment that they can't do now for lack of information?[43]

Arguably, the most important step in an assessment is specifying the players: Who will be the focus of the study; who requested the study; who will gather and analyze the data; and who will use the information gathered during the study. The objects of the study are the participating students and instructors. Realistically, whoever has requested the study should want to use the results, even though other people or agencies may benefit. Since they are the primary intended users of the assessment, the individuals who requested the study must clarify its purpose.[44] The presence of an identifiable individual or group of people who personally care about the assessment and the information it generates is crucial, as the absence of a personal factor severely diminishes the impact and the use of the assessment.[45] Similarly, it is crucial to involve, from the beginning, those people who are responsible for gathering and analyzing the collected data in developing the assessment methodology; their contributions enhance a successful assessment. Educators also must include the instructors participating in the study in developing the assessment methodology, as they will develop and teach the course that the educators will assess.

The individual or group of people responsible for the assessment also must focus the assessment on a specific goal or set of goals. This focusing insures that the assessment concentrates on those issues which will produce the most useful information, yield the greatest insights, and provide the most solid basis for action.[46] In addition, by specifying assessment goals, educators can frame the assessments so the results address relevant issues. The educators can begin to formulate these goals by answering such questions as:

1. What is the purpose of the assessment?

2. Is the methodology intended to assess teaching effectiveness based on a cost-benefit analysis?

3. Is the time required for teaching and/or learning an issue?

4. Should the assessment determine if students have attained specific objectives at a desired level of learning?

5. What comparisons of the experimental and control groups should educators attempt?

6. What information is needed and wanted by the people who intend on using the assessment?

7. Whose values are educators to apply in the assessment and in making subsequent judgments and evaluations?

8. Does short-term assessment of objectives reflect actual learning or would educators find long-term student tracking more revealing?

9. Given the course to be examined, is the assessment realistic and can educators accomplish it?

10. What are the temporal restrictions on the assessment with respect to conducting only one study or many studies spread over several years?[47]

Clearly, educators need to address many other questions to properly focus the assessment. The process of developing the assessment goals must be interactive, involving relevant decision makers, information users, instructors, and evaluators.[48]

After the people responsible for the assessment have specified their goals, educators can identify or develop the appropriate assessment methodologies. To compare the effectiveness of two different teaching styles, educators must examine the performances of the students exposed to those styles. Hence, educators need to clearly identify and define the specific performance parameters of interest to insure that they can indeed collect the required data. They must then find or develop methodologies to measure these parameters and judge the appropriateness of the methodologies according to the goals of the assessment.[49]

With the methodologies in place to measure student performance, educators can then generate and analyze the data. A final, but critical, step in any assessment process assembles the results so educators can use them effectively. The decision makers must specify how they plan to use the information obtained from the assessment to make judgments, evaluations, and decisions so that evaluators know what to analyze and present.

Summary

Assessing the effectiveness of two different teaching styles requires a comprehensive assessment strategy. Educators must identify and relate the goals of the assessment to appropriate assessment methodologies. These methodologies, coupled with a well-defined baseline scenario which characterizes the course, students, and instructors, are prerequisites to a successful study of teaching effectiveness. If the people involved with the study thoroughly address the issues discussed in this chapter, then they can accomplish a study to assess the effectiveness of different teaching styles.

However, if educators wish to examine the effectiveness of innovative educational technologies as teaching aids, they can only accomplish a valid appraisal of the effects of these technologies on student performance by developing appropriate assessment methodologies which can clearly distinguish between the influence of educational technology and different pedagogies.[50] Similarly, educators assessing the effectiveness of novel teaching styles that are still traditional approaches to teaching must develop methodologies which can discriminate between the pedagogies and other influences on student learning.

Before educators can accomplish any study of teaching styles, they must first understand the various aspects of the different teaching styles. Chapter 2 briefly reviews many such teaching styles by describing traditional teaching methods and teaching using innovative educational technologies.

Notes

1. *Webster's II New Riverside University Dictionary* (Boston: The Riverside Publishing Company, 1988), 1187.
2. Andrew V. Beale, "Are Your Students Learning What You Think You're Teaching?," *Adult Learning* 4, no. 3 (January/February 1993): 18.
3. Dr Kathleen Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), interview with author, 19 January 1993.
4. Ibid.
5. Beale, 18.
6. Dr Michael Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), interview with author, 19 January 1993.
7. Dr Kathleen Hannafin, interview with author, 13 August 1992.
8. Dr Mary Marlino, visiting professor, Directorate of Education, Air Force Academy (HQ USAFA/DFER), interview with author, 14 August 1992.
9. Ibid.
10. Lt Col John Souders, director of research, Directorate of Education, Air Force Academy (HQ USAFA/DFER), interview with author, 11 August 1992.
11. Dan Surry, graduate assistant, Directorate of Education, Air Force Academy (HQ USAFA/DFER), interview with author, 11 August 1992.
12. Richard J. Shavelson, Noreen M. Webb, and Leigh Burstein, "Measurement of Teaching," in *Handbook of Research on Teaching*, 3d ed., ed. Merlin C. Wittrock (New York: Macmillan Publishing Co., 1986), 50–91.
13. Ibid.
14. *Webster's II New Riverside University Dictionary*, 418.
15. Richard E. Clark and Gavriel Salomon, "Media in Teaching," in *Handbook of Research on Teaching*, 466.
16. Dr Kathleen Hannafin interview, 13 August 1992.
17. B. S. Bloom, *Taxonomy of Educational Objectives: Book 1, Cognitive Domain* (New York: Longman, 1957).
18. Dr Kathleen Hannafin interview, 13 August 1992.
19. Michael Q. Patton, *Utilization-Focused Evaluation* (Beverly Hills, Calif.: Sage Publications, 1978), 133.
20. Dr Kathleen Hannafin interview, 19 January 1993.
21. John B. Carroll, "The Carroll Model, A 25-Year Retrospective and Prospective View," *Educational Researcher* 18, no. 1 (January/February 1989): 30.
22. Ibid.

23. Nancy S. Cole and Anthony J. Nitko, "Measuring Program Effects," in *Educational Evaluation Methodology: The State of the Art, ed. Ronald A. Berk (Baltimore: Johns Hopkins University Press, 1981), 32–63.*

24. Anthony S. Bryk and Richard J. Light, "Designing Evaluations for Different Program Environments," in *Educational Evaluation Methodology*, 17.

25. Ibid., 23.

26. John Biggs, "Teaching for Desired Learning Outcomes," in *Handbook of Educational Ideas and Practices*, ed. Noel Entwistle (London: Routledge, 1989), 682.

27. Ibid.

28. Surry interview.

29. Cole and Nitko, 44.

30. Noel Entwistle, "Teaching and the Quality of Learning in Higher Education," in *Handbook of Educational Ideas and Practices*, 669.

31. Ibid., 672.

32. Ibid.

33. Ibid., 673–74.

34. Lt Col William Newmiller, Department of English, Air Force Academy (HQ USAFA/DFENG), interview with author, 11 August 1992.

35. Cole and Nitko, 45–50.

36. Shavelson, Webb, and Burstein, 58.

37. Robert B. Kozma and Jerome Johnston, "The Technological Revolution Comes to the Classroom," *Change* 23, no. 1 (January/February 1991): 19; Maj Judith Brisbois, Department of Foreign Languages, Air Force Academy (HQ USAFA/DFF), interviews with author, 12 August 1992 and 19 January 1993.

38. Ibid.

39. Dr Michael Hannafin interview, 12 August 1992.

40. Dr Michael Hannafin interview, 19 January 1993.

41. Marc J. Rosenberg, "Evaluating Training Programs for Decision Making," in *Evaluating Business and Industry Training*, ed. Leslie S. May, Carol A. Moore, Stephen J. Zammet (Boston: Kluwer Academic Publishers, 1987), 58; Irwin L. Goldstein, "Content Validity as an Evaluation Strategy for Examining Training Programs," in *Evaluating Business and Industry Training*, 139.

42. Michael Q. Patton, *Creative Evaluation* (Beverly Hills, Calif.: Sage Publications, 1981), 23.

43. Ibid., 167; Michael Q. Patton, *Utilization-Focused Evaluation* (Beverly Hills, Calif.: Sage Publications, 1978), 81–82.

44. Patton, *Creative Evaluation*, 145.

45. Patton, *Utilization-Focused Evaluation*, 64.

46. Patton, *Creative Evaluation*, 149.

47. Ibid., 186–92.

48. Patton, *Utilization-Focused Evaluation*, 81.

49. Cole and Nitko, 34.

50. Dr Michael Hannafin interview, 12 August 1992.

Chapter 2

The Learning Environment:
Traditional Versus Educational Technologies

What is the purpose of teaching? Quite simply, teaching imparts knowledge or skill and advocates or causes learning by example or experience.[1] Teaching does not occur in a vacuum but within the context of a learning environment. Educators purposefully create a learning environment to impart knowledge and understanding or to inspire learning. Michael Hannafin contends that "learning environments share a variety of dimensions: the scope of the environment, learner and educational activities, and content integration. Each dimension exists as a continuum, and learning environments possess attributes along each continuum."[2]

Scope refers to the different characteristics of the environment and includes such aspects as the information covered and the educational tools that are available to the learner. The continuum for this dimension ranges from the micro to the macro level. On the one hand, micro-level learning environments provide a high degree of focus within a subject area to examine, explore, and synthesize, in detail, interrelated information, skills, and concepts. On the other hand, macro-level learning environments broaden the coverage of the course content, relaxing the emphasis of focus exhibited in the micro-level environment and enabling students to pursue external, but related, interests or needs.[3]

The nature of learner activity adds another dimension to the learning environment. At one end of the spectrum, the environment relies "on the individual (or group of individuals) to create, elaborate, or otherwise represent knowledge."[4] Environments at the opposite end of the spectrum "vary the manner in which information is organized as well as the method in which it is provided";[5] content is structured "to permit the student to learn according to externally generated notions of meaning."[6] In other words, the continuum of learner activity ranges from student-centered to teacher-centered learning.

Learning environments also vary as a function of the nature of the educational activity. In this case the spectrum of the environment varies from goal-directed, intentional learning to student-directed exploration. As Michael Hannafin states,

> Goal-directed environments emphasize intended competence, facility, or comprehension. The activities are designed to support a defined set of learning outcomes. . . . Exploratory environments emphasize processes more than outcomes least insofar as intended learning is concerned. Often, students are encouraged to alter, explore, or otherwise manipulate the parameters of the environment to examine possible outcomes.[7]

The final dimension to consider in the learning environment focuses on content integration. In many cases, the environment concentrates on the subject and integrates closely related skills and concepts within the limited context of the course. At the other end of the spectrum, the learning environment minimizes the explicit or implicit boundaries associated with a given topic or subject to integrate into the course information, concepts, and skills outside those boundaries.[8]

Teaching Models

Maximizing the performance of students in classrooms requires that educators gear their styles of teaching to specific points on the continuum of each dimension within the learning environment. Educators usually base these styles on a set of teaching models. Bruce Joyce and Marsha Weil have identified four fundamental models of teaching to provide a basis for developing teaching styles: social interaction; information-processing; personal; and behavior modification.[9]

The first teaching model is oriented towards social relations and emphasizes the relationship of people to society or other people. The social interaction model improves an individual's ability to relate to others. Some styles of teaching which use this model specifically improve social relations, while others use social relationships as the primary vehicle of education.[10]

The second, the information-processing model, develops or improves a student's capability to interpret and to process information. Associated teaching styles concentrate on expanding the human capacity for processing information, increasing the student's ability to handle stimuli from the learning environment, to organize data, to understand problems, to generate concepts and solutions to those problems, and to employ appropriate verbal and nonverbal symbols when solving the problems.[11]

The third, the personality teaching model, develops a student's personality as a source of educational ideas. Teaching styles therefore concentrate on personal development and emphasize processes by which an individual constructs and organizes his or her reality as a means of acquiring and processing knowledge, skills, and higher order learning abilities and understanding.[12]

The fourth teaching model is developed from analyzing the processes by which human behavior is shaped and structured. Styles implementing this behavior modification model create efficient systems for sequencing learning activities and shaping behavior by manipulating how instructors reinforce a student's learning process; in essence, these styles attempt to change the external behavior of the student to enhance his or her ability to learn.[13]

Teaching styles developed from one of the four models outlined above have one common goal: to increase a student's performance in a classroom environment. The means by which the styles educators can use to accomplish

this goal vary greatly, as evidenced by the different teaching models; however, they must realize that effective teaching styles normally integrate aspects of all four models. Consequently, the educators should understand the various styles they can use to teach students. For this study, I have classified teaching styles into two broad categories: traditional teaching methods and teaching using innovative educational technologies. This chapter examines the more common teaching styles in each category and highlights some of their advantages and disadvantages.

Traditional Teaching Methods

In many ways, traditional teaching methods are associated with the "old ways" of conveying knowledge to students: memorization, drill, essays, phonics, rules of mathematics, diagramming sentences, and basic concepts of citizenship, ethics, geography, discipline, and individual, family, and community responsibility.[14] Frederick Hill states that "the old ways showed, on the blackboard, how you reasoned and arrived at an answer. The new ways ignore the skills the old ways generated."[15] If the "new ways" of teaching imply using innovative educational technologies as teaching aids, which teaching styles promote the old ways or traditional teaching methods?

Didactic Teaching

The most common style of traditional teaching is catechetical, or didactic, teaching, which is conceived as an expert imparting knowledge or explaining or describing phenomena to a group of learners.[16] This teaching style invites the student to acquire a body of knowledge as defined and presented by an expert. It is simple, systematic, and usually passive: students speak only in response to questioning by the expert or instructor. Lectures are familiar didactic teaching styles and are the most efficient and economical way to transmit a common body of knowledge to a large number of students.[17]

One can argue that lectures are unnecessarily formal and authoritarian, promoting passivity among students.[18] Many academicians, in fact, are convinced that lectures are ineffective and that other teaching styles facilitate the learning process by involving the students and making them active agents in developing their own knowledge, skills, and understanding of the topics they are learning.[19] These educators contend that structured lectures only limit the student's autonomy and, therefore, ability to learn in the classroom.[20] For example, E. Peter Volpe believes that society's ignorance of the science disciplines can be traced to how educators teach science:

> The inability of students to appreciate the scope, meaning, and limitations of science reflects our conventional lecture-oriented curriculum with its emphasis on passive learning. The student's traditional role is that of a passive note-taker and regurgitator of factual information. What is urgently needed is an educational program in which the students become interested in actively knowing, rather than passively believing.[21]

However, as Jim Corder states, "Arguments that lectures are pedagogically unsound are invalid because they are based only on bad lectures. We don't often examine what makes a lecture good or differentiate the good from the bad, so we wind up opposing all lectures."[22] Many educators believe that instructors supposedly design lectures to disseminate knowledge; if this were the case, there would be better ways to pass knowledge on to students: library searches, reading assignments, and the like. In actuality, lectures dramatize knowledge and offer an interpretation of that knowledge on to students. No single style of teaching guarantees student learning, nor is a particular style right for everyone in all circumstances. Disregarding the lecture as a valid teaching style ignores differences in personalities and preferences for teaching and learning among teachers and students.[23]

Mastery of content is one distinct advantage of a lecture format. Lectures are also the students' preferred means of preparing for exams; the information disseminated during the lectures is usually what is tested on the exams. At the same time, however, even though lectures increase the focus of a student's attention, they reduce his or her appreciation and understanding of peripheral issues. Lectures not only inhibit discussions, which could enrich the lesson, but may also increase the potential for a student's boredom.[24] In fact, structured lectures are "vulnerable to routinization of learning and a concomitant dehumanization of students."[25] Even though the structured lecture provides a positive, but rigid, supporting format for the instructor in preparing the material for class, too much structure limits his or her creativity and spontaneity.[26] The instructor may concentrate on presenting a set amount of material within a specified time and fail to regard each student as an individual who may require more personal interactions to understand the material.

Didactic styles of teaching are not limited to lectures. Rather, they include a wide range of teaching styles which promote active learning yet do not primarily rely on innovative educational technologies as teaching aids. For example, relaxing the structured format of lectures introduces various other teaching styles, such as inquiry teaching or cooperative teaching.

Inquiry Teaching

Inquiry teaching requires the students to be insightful while they are exposed to a lecture-based, classroom environment. One form of inquiry teaching is the "think-aloud modelling" approach, where the instructor thinks out loud to the class while performing a complex activity.[27] Students learn how the instructor develops ideas or solutions to problems, which helps them to learn how to formulate their own thought processes.

The student self-questioning technique, another style of inquiry teaching, requires each student to pose questions periodically to him- or herself regarding the material being presented in class; thus, the student can monitor his or her comprehension during the class.[28] If the student doesn't understand the material, he or she can ask the instructor for clarification.

Cooperative Teaching

In cooperative teaching, an extension of the lecture style of teaching, a team of instructors teach the students. This didactic teaching style addresses the misconception that the teacher is the established expert in all aspects of a subject area. In reality, the instructor's expertise is usually limited to a narrow range of topics.

The main aim in cooperative teaching is to match the demands of the curriculum with the range of expertise of a team of instructors. Teachers with different backgrounds and experiences complement each other in class by bringing in their own perspectives and experiences to the course content, approach, assessments, resources, and organization to enhance the quality of classroom teaching and learning.[29]

Student-Centered Learning

As implied in the paragraphs above, didactic teaching styles center the learning around the instructor. Students are usually in a passive learning mode, looking to the instructor for guidance and to impart knowledge and understanding. These instructor-centered teaching styles are usually organized for groups of learners; the instructor makes key decisions for these learners and provides them with preset conditions under which they must perform. Learning can also have a student-centered orientation. However, in this environment the student takes an active role in the learning process by shouldering the responsibility for learning the required knowledge, skills, and higher order learning abilities and understanding. Michael Hannafin states that,

> Student-centered learning systems essentially define the student as the principal arbiter in making judgments as to what, when, and how learning will occur. [These] learning systems tacitly presume that students possess the metacognitive skills needed to make effective judgments, or that they can be induced to make appropriate choices through advice, hints, or guided reflection.[30]

Successful student-centered styles of teaching offer a sufficient array of resources for students to use as their knowledge, skills, and abilities evolve. The instructor provides support and encouragement to the students rather than solely transmit knowledge and information to them.[31] Traditional teaching methods incorporate a variety of such student-centered teaching styles as guided discussions, seminars, role-playing, syndicate teaching, and open learning.

Discussions

Instructors can use discussions in conjunction with lectures or by themselves as mediums of learning. The instructor promotes discussions by first making statements or asking questions and then inducing student responses. The central feature of the discussion teaching style is the opportunity for students to learn through their expressing and exploring diverse ideas and experiences in a cooperative learning environment. Its

emphasis is not on winning arguments but on using the diverse resources and expertise available in a group to enhance the students' understanding of a subject, sharpen their judgments, and extend their knowledge.[32]

Instructors can apply the discussion style of teaching in the constructive classroom, where teaching is pursued through conversation. This environment allows teachers and students to discuss ideas and concepts and to negotiate, through conversations, to construct shared meanings of these ideas and concepts. The instructor still serves as the expert in the subject matter, and this teaching style emphasizes the quality of interaction as students increase their knowledge of the subject. The teacher asks probing questions to determine if the students have grasped particular concepts and to discover their misconceptions. A critical requirement of this style of teaching is to establish specific objectives for the course so teachers have performance measures when they use the conversations to assess student understanding of the concepts.[33]

Seminars

The seminar approach to teaching closely resembles the discussion teaching style, although it usually has a smaller group of students than the discussion approach, and the instructor designates one person as the seminar leader. While the instructor remains involved in discussions as a neutral expert and is still the focal point of the class, the seminar leader facilitates interactions among the students yet, at the same time, remains as an outside observer. The instructor is considered the expert in discussions; this arrangement is not necessarily the case for the seminar leader. Seminars help students define the individuality of their own thought processes in addition to learning the discourse of the subject. Students understand and discuss not only the subject knowledge but also related concepts to gain further understanding of the subject.[34]

Teachers can frequently use discussions and seminars more effectively than didactic styles of teaching to stimulate students to think, to comprehend, or to examine their attitudes. In addition, these two teaching styles can enhance a student's personal adjustments to and interactions in social situations while simultaneously transmitting information almost as effectively as didactic styles. Using discussions and seminars as styles of teaching requires a great deal of constraint from the instructor, who acts as a neutral chairperson or facilitator; as such, the instructor doesn't endorse any single perspective on an issue. The instructor encourages the students to search critically for meanings in the subject matter. The instructor must be well organized, a good listener, and an exact questioner.[35]

Study Circles

Students do not always require instructors to guide their learning; a study circle is a collaborative style of learning which needs no active instructor involvement. An abiding tenant of a study circle is voluntary and equal democratic participation; every member in the group has an equal voice, and

participants determine the agenda and outcomes. Ideally composed of between five and 20 members, a study circle is a small group setting which encourages individuals to function democratically and to learn individual and collective decision-making skills and abilities while developing self-directed learning abilities.[36]

A study circle requires no content experts or, for that matter, instructors, but has a well-prepared leader who maintains control and encourages participation. Members use print, video, and audio materials to stimulate discussions, frame the session, and establish a logical sequence for the discussions. The leader can use a guide which contains sample discussion questions, ideas related to dealing with group dynamics, questionnaires, and goals and objectives for the session. Since learning is collaborative rather than competitive, study circle participants learn from cooperation while learning the content and objectives of the session. Participants work together to understand material, analyze problems jointly, and arrive at common conclusions. Study circles are most effective with social, political, economic, and cultural disciplines instead of subjects requiring individual mastery of complex principles and masses of information. The participants must be mature enough to function in such an open environment, and educators who use study circles must anticipate that the participants may not achieve the course goals without close supervision.[37]

Role Playing

Instructors must be innovative to entice students to participate in discussions, seminars, and study circles. Another novel style of teaching educators use to facilitate active student involvement in the classroom is role playing. Role-playing exercises provide the student with the opportunity to empathize with the actions of another person to help understand how that person thinks and feels. One form of role playing in, for example, a discussion or debate occurs when the teacher assigns students such interconnected roles as presenting information, observing other participants' responses and behaviors, and mediating the discussions of different points of view.[38] This form of role playing helps the students to learn how to recognize and manipulate information while they develop practical learning skills and abilities. Other role-playing techniques task students to act out the role of a particular person within the constraints of the personal characteristics and expectations of that individual, assumptions and expectations of others, and the physical limits imposed by the environment.[39]

Games and simulations are applications of role playing. In a game, one or more players compete, or cooperate, for payoffs according to an agreed upon set of rules. Students succeed by properly using the information at hand. While games are usually highly structured, they don't have to model reality. A simulation possesses the same features of the game but is a dynamic, working representation of a real-life situation. It may be abstract, simplified,

or accelerated, yet a simulation incorporates critical features of reality to make it believable to the students.[40]

A key advantage of these three styles of teaching—role-playing, games, and simulations—is the increase in student involvement in the classroom environment. Students embrace an active learning process and are more apt to inquire and delve into the subject in greater detail than with passive teaching styles. These active styles of teaching may lead to long-term retention of the material and to increased understanding of the issues surrounding the subject. Finally, there is usually a high level of enjoyment and motivation among the students.[41]

As for the instructor, he or she is considered the neutral chairperson or facilitator of the exercises rather than the source of knowledge. This role requires special skills of the instructor, and the success or failure of the class depends heavily on the quality of that instructor. A major concern of instructors in using these teaching styles is the depth of content coverage. These styles require a great deal of time from the students as they delve into the subject; as such, the students may not adequately cover the required material. In addition, the styles have a tremendous appetite for the time and resources of the instructors. They must carefully design the lesson, as simplifications can potentially mislead the students. As facilitators, the instructors must be aware that erroneous assumptions and information the students contribute are sometimes subconsciously incorporated into the learning process. There is always the danger of hidden agendas promoted by either the instructors or the students. Finally, instructors must periodically debrief the students to clarify issues and insure that students take away from class something of value; in fact, the instructors should always attempt to extrapolate the results of a majority of the lessons to the real world.[42]

Syndicate Teaching

While role-playing styles of teaching generate student involvement by immersing them into situations, syndicate teaching—also referred to as cooperative learning—uses teamwork to involve students in an active learning environment. Small teams of between four and eight students explore the various aspects of a problem or topic the instructor assigns. Drawing on appropriate sources and investigating techniques, each group summarizes its findings in a comprehensive report. Syndicate teaching develops a student's powers of critical judgment by applying learned material in entirely new settings, solving problems in unfamiliar ways, analyzing the structure of a given argument or communication, exercising inventive or creative talents, judiciously assessing the quality of given materials, and expressing ideas or feelings in a variety of media.[43]

In syndicate teaching, students enhance their capability to learn when they participate in group settings rather than trying to learn on their own or in competition with others. In group settings, students learn how to cooperate to complete an instructor's tasking; each student's success depends on and is

linked to the success of every other student, and the outcome of the tasking consequently results from the students' common effort.[44]

Syndicate teaching methods promote teamwork, as individuals work toward a common goal, and mutual dependency motivates them to work hard to help the group, and themselves, succeed. This method also accelerates a student's intellectual development by forcing the student to assess group discussions systemically and confront other students who hold opposing points of view. Ideally, the group works together to achieve a common goal; this process necessarily leads to collaboration, either through debate, argument, negotiation, or discussion. Compromise is indispensable for cognitive growth. Finally, by tutoring others, students learn more when they help others in the group; students also may learn more when others help them.[45]

To apply syndicate teaching successfully, instructors must continually monitor each group's progress to insure that students remain actively involved in their project. Syndicate teaching is effective if the students develop not only a strong cooperative involvement in academic inquiry but also their higher order learning abilities and understanding. Indicators of involvement include better class attendance, a willingness to invest time and energy in the group, a stronger commitment to work outside of the class, a stronger sense of mutual obligations, and an effort to actively search for information. Syndicate teaching allows students to demonstrate a more critical approach to reading and research and to appreciate more fully a variety of opinions and experiences.[46]

While syndicate teaching strives to develop a student's higher order learning abilities and understanding, it assumes the student already reasonably comprehends the subject. The benefits associated with this teaching style diminish if the students are unfamiliar with the subject. Students may need to conduct extensive background research on particular topics before they can contribute to the group.

Implementing syndicate teaching in the classroom requires a tremendous amount of preparation time from the instructor to match subject matter to the subject comprehension level of the students, as well as to match individual students to insure cohesive working groups. Instructors must be sensitive to the student who becomes passive in such uncomfortable situations as working groups. Finally, groupthink or dominant personalities may influence work within the group, requiring the instructor to monitor closely the group to prevent undue bias and to keep it on track.

Problem-Based Learning

A variation of syndicate teaching, problem-based learning provides educators with a vehicle to make students active participants as they learn how to solve problems. Typically, students acquire knowledge that relates to resolving a problem; the body of factual knowledge becomes meaningful when students use it to solve a problem. Problem-based learning presents a student with a problem or situation.[47] Then,

> he or she must seek out the relevant information, formulate a hypothesis consistent with the observations, deduce predictions from the hypothesis, and subsequently accept, reject, or modify the hypothesis in accordance with the degree of fulfillment of the predictions.[48]

This teaching style allows the student to develop those inquiry skills and abilities that are characterized by objectivity, open-mindedness, and skepticism. It is most effective when the instructor organizes the students in small discussion groups and encourages them to actively participate in the learning process. The members of the group benefit from their diverse backgrounds and levels of knowledge and understanding by learning from and with each other as they work together to solve a problem. The instructor, while an expert in the discipline the students are studying, assumes the role of a facilitator and therefore must be skilled at guiding discussions and assembling appropriate learning materials that are relevant to the problem at hand. The students retain the initiative in problem-based learning and must establish effective working relationships if they are to successfully solve the problem the instructor has posed. As the students meet the instructor's explicit goals, they highlight their strengths, weaknesses, attitudes, and prejudices, and they learn to become more inquisitive and open-minded.[49] Unfortunately, the same drawbacks of syndicate teaching apply to problem-based learning.

Open Learning

The previously mentioned instructor- and student-centered styles of teaching do not provide the students with the opportunity to negotiate the course objectives nor their assessments. Open learning is a student-centered teaching style that is flexibly designed to meet individual requirements. Open-learning classes enable the learners and the instructor to negotiate the course content as well as the learning and assessment methodologies. Open learning develops the student's skills of negotiation and communication and promotes self-motivation.[50]

Instructors run most courses that use traditional teaching methods at a set time and place; conversely, the resources for open-learning courses are available anytime and anyplace. This availability allows learning to be accomplished at different rates, so the students are self-pacing and can choose their preferred sequences of learning subject matter. However, not all open learning takes place alone or with an instructor or tutor; students may be brought together for such group activities as lectures, discussions, or role playing.[51]

Although open learning encourages the students to develop their higher order learning abilities and understanding, it isolates the students unless they are involved in group activities. Open learning also requires a high level of motivation and discipline from each student, and students can become frustrated easily. Hence, the instructor must be imaginative in monitoring student progress to maintain a balance between tracking his or her progress while not overcontrolling the student.[52]

Teaching Using Educational Technologies

As previously mentioned, educators claim that student-centered teaching styles are advantageous to didactic teaching in that learning is student-centered as opposed to the instructor-centered approach of the didactic teaching styles. Still, it is insightful to mention an assertion made by Jim Corder:

> Yet I don't believe we can assume that flashes of insight and the illumination of new knowledge will occur any more readily in tutorials, small group discussions, student-led classes, or any of the other methods of learning now in favor than they would if an instructor were dominating the conversation. . . . Mostly, epiphanies have to be earned, which happens best through preparation and hard work by both students and faculty members—not through the random insights of haphazard student contributions.[53]

Many professional educators would modify Corder's assertion as a reflection of their opinion of innovative educational technologies. Some people believe that educational technology is "education's biggest boondoggle ever."[54] In contrast, proponents of educational technology claim it is the "best hope of educating an increasingly diverse population."[55] What, then, is educational technology, and why does it create such a disparity of views?

According to Dr Michael Hannafin, educational technologies are those systems through which media and teaching styles interact to induce learning; they focus on the system rather than on the hardware. Educators use technologies, or media, in conjunction with classroom activities to convey knowledge. The existence of educational technologies in a learning environment does not preclude the instructor from teaching. In fact, they enhance the instructor's role by allowing the instructor to interact with students to insure that learning occurs.[56]

To be sure, educational technologies have been around for a long time. However, their use has usually been limited to such basic audiovisual aids as slide and motion picture projectors, chalkboards, and overhead projectors. Usually, instructors use these educational technologies in the classroom to aid their traditional teaching methods. Still, according to Ken McCullough, "The instructor-centered, instructor-dominated, workbook-oriented class in most [educational] programs is little different from that which was present several decades ago.possess."[57] There are many legitimate reasons for the lack of use of educational technology; some include the high costs of purchasing the required hardware and software, inadequate funds for maintenance, service, upgrades, and licensing fees, inappropriate or inadequate software and hardware, the absence of assessments to examine the complex thinking skills and abilities the technology rich environments encourage, and high rates of technological illiteracy among university faculties.[58]

In recent years, though, innovative educational technologies have "launched a revolution in learning and teaching in higher education."[59] The popularity and use of innovative educational technologies have increased due to their expanding capabilities coupled with decreasing hardware and

software costs. Multimedia presentations may enhance lessons by stimulating student creativity and motivation. Further, the student's function in the educational process can be more of an active participant. Technology can make learning more exciting for students and can reduce the possibility of boredom, which often causes low test scores and substandard performance in class; however, boredom doesn't necessarily cause students to get poor test scores.[60] The approach to teaching is fundamentally different when an instructor uses innovative educational technologies; he or she can change qualitatively the nature of the learning experience itself by using a variety of learning activities.[61] For example, rather than simply reciting lessons, the instructor can interact with individual students during class; while some students use computer resources to explore a topic, others work individually or with assistance from the instructor.[62] The instructor does not need to be an expert in every topic; innovative educational technologies reduce the instructor's need for specific knowledge and expertise and provides the opportunity not only to guide the students through the learning process but also to learn along with them.[63]

Qualitatively changing the approach to teaching requires a catalyst: many educators claim that innovative educational technology is that catalyst. Numerous technologies exist which could support active learning and adventurous teaching; in fact, some of the older technologies, or media, have enriched the learning environment. For example, film strips and movies provide visual examples of real-world phenomena and events to bring reality into the lesson. Instructors can use overhead projectors to show slides or clarify complex concepts in much the same, albeit neater, way as an instructor who uses a simple chalkboard. Some instructors, however, have used such technologies as a crutch for poor teaching; others are just lazy and have relied on the technologies as substitute teachers. Thankfully, these instructors are a minority; most use educational technologies to enhance their lessons.

While educational technologies traditionally have included film strips, slide shows, movies, overhead projectors, and chalkboards, current innovative technologies include more exotic media: computers, computer networking, satellite communications, videodiscs, and multimedia technologies. Understanding the potential of these innovative educational technologies necessitates a brief description of some of them and their applications. The following discussion of innovative educational technologies is by no means comprehensive but provides a sampling of existing applications to clarify the various types of innovative educational technologies that are available.

Videodisc Technology

Videodisc technology is simply an extension of film strips, slide shows, and movies. It bridges the gap between the classroom environment and the real world by immersing students in real-world situations in which they exercise problem-finding and problem-solving activities.[64] Robert Kozma and Jerome Johnston believe that "too often our students walk out of class ill-equipped to

apply their new knowledge to real-world situations and contexts. Conversely, too frequently the classroom examines ideas out of the context of gritty real-world considerations."[65]

Applications of videodisc use in a learning environment are abundant; mentioning only a few demonstrates the diversity of videodisc technology. For example, one interactive video program provides students of ethics and medicine with a situation for testing their ethical principles and judgments by presenting an actual case of a serious burn victim who sees death as a way to end his pain and suffering. This graphic video presents students with conflicting arguments from some of the people involved in the case, and students must formulate a reasoned response to the patient's request.[66]

Videodisc technology has also made its way into scientific laboratories. As Kozma and Johnston argue,

> The scientific laboratory is one of the most expensive arenas in the [academic environment]. It is costly to maintain the proper equipment and supplies, and to provide supervision to student scientists. It is also limited as a learning experience. So much time is required to replicate classic experiments in the discipline that there is little time left for students to explore alternative hypotheses as real scientists do.[67]

"Chemical Reactions" is a scientific videodisc series which shows 14 chemistry experiments that are expensive and dangerous to conduct in a real laboratory. Students interact with the video by choosing a combination of equipment and strategies and witness the resulting experiment in an authentic video sequence. This program fosters the investigatory skill of chemistry students; the students actively participate in the laboratory even though they are not actually mixing the chemicals.[68] In addition, they can make mistakes and not waste valuable resources or endanger themselves. Finally, they can develop their investigative abilities by choosing alternative approaches to the experiments and predict and observe the results.

The Department of Foreign Languages at the Air Force Academy has used an interactive videodisc system for several years. For example, students use the "French in Action" video series in their language study. They witness authentic cultural scenes and speakers and become immersed in situations designed to introduce and reinforce various aspects of the language. The video system is interactive in that the students mentally participate in the storyline and can verbally respond to questions from participants in the video. They become involved in the storyline and try to anticipate the next occurrence. In addition, the myriad of exercises permits the student to move around within the program, and they can repeat segments of the program or specific phrases until they are confident of their understanding. Thus, the students proceed through the program at their own pace.[69]

By controlling the pace of learning, students can move through the material as quickly or as slowly as they desire, and they can spend extra time on topics which they find difficult. However, this self-pacing is not effective in all situations; students must be able to understand the relationship between spending time on a topic and actually mastering the topic. Often, instructors

must coach them on the self-paced study, a technique which can strengthen both their performance and motivation to learn by themselves.[70]

The French instructors at the Air Force Academy promote efficient self-pacing and achieve an effective learning environment by complementing the "French in Action" videodisc program with in-class practice. After the students complete a lesson on the videodisc, they meet in class to practice the language and concepts presented in the program. By interacting and communicating with each other, the students reinforce the concepts and listening skills they are developing through the videodisc program; in addition, they can practice and develop their oral communication skills when they speak French in the classroom.[71]

Computer Software Packages

While the above examples illustrate how videodiscs introduce a variety of interactive applications to students in a learning environment, computers provide an even wider variety of applications. Computers introduce information, ideas, and examples into the classroom that many educators never before thought possible. According to Allan Collins,

> computer technology can be used in the classroom in three ways: as tools such as word processors, spreadsheets, programming languages, and electronic network systems; as integrated learning systems that present exercises for students to work on individually and keep records of student progress for reporting to the teacher; and as simulations and games which engage students in computer-based activities designed to motivate and educate.[72]

Early uses of the computer were limited primarily to drill and practice, tutorials, and computer-assisted instruction. Certainly an advantage of the computer involves its capacity to store and manipulate tremendous data bases on any topic of interest. Drill and practice programs do not present new information to students; rather, they test existing knowledge in a variety of ways. In many cases, the computer monitors the student's progress, analyzes strengths and weaknesses, and adjusts the lesson accordingly. The results can be compiled and aggregated with the rest of the class to provide concise reports to the instructor.[73]

As a tutor, the computer teaches specific facts by modelling the skills instructors expect students to learn. As such, James Kulik and Robert Bangert-Drowns believe educators design computer tutors to:

1. gain the student's attention
2. stimulate the recall of prerequisite information
3. state the lesson objectives
4. present relevant stimuli
5. provide guidance
6. elicit student performance
7. assess the performance and provide feedback
8. ensure the retention and transfer of the newly learned information.[74]

Instructors can use computer tutorials as cognitive tools to provide students the resources and opportunities to teach themselves. Computer-assisted teaching also allows students to teach themselves, but in this case the computer goes beyond tutoring by forcing the students to learn knowledge beyond just mere facts. The computer contains large data bases and complex algorithms, and students interact with the computer to solve problems and make predictions and reasonable guesses about situations. Finally, the computer provides feedback on the students' work and offers advice to the students about their efforts.[75]

Amy Rose espouses several benefits from using computers in educational programs.

> They allow for greater individualization in the learning process, provide immediate feedback, . . . thereby further enriching the learning process. Because they are working alone on a computer, the students have a greater control over their own pace; they can work in a risk-free environment, and at their best, computers allow students to express their curiosity and demonstrate their creativity. . . . One of the most highly touted aspects of computers . . . is the amount of control the learner gains over the education process when computers are used . . . such control can raise self-esteem . . . it also presents the possibility of following interests on an individual basis.[76]

Computer-assisted teaching also benefits the instructors. Instructors can relegate basic drills and other behavioral-related teaching to the computer and concentrate on teaching higher order skills, abilities, and understanding. However, critics claim that this methodology dehumanizes the learning process and increases the social isolation of the learner. They also believe that those educational programs which use computer-assisted teaching discourage dialogue and interaction; consequently, student growth and change cannot occur.[77]

Computer-assisted teaching is just one example of how computers can enhance a student's learning environment. The capabilities of computer simulations resemble those exhibited by videodisc technology in that they allow students to accomplish tasks not feasible in ordinary classroom environments. However, while videodiscs immerse students in real-world situations, computer simulations permit students to visualize more abstract concepts. For example, physics students can use simulations to observe and modify their misconceptions of physics phenomena and to develop an understanding of physics principles. They can simultaneously measure and graph such physical quantities as position, velocity, acceleration, and force. The students can then conceptualize these variables in terms of physical phenomena they can understand.[78]

Similarly, students can use computer simulations of realistic problems to generalize the results to naturally occurring phenomena. Students often find it difficult to generalize the knowledge and understanding they acquire in science courses to the real world; they cannot relate their classroom experiences to complex and ambiguous problems outside class nor can they apply abstract models of scientific phenomena to concrete real-world

occurrences. Students can gain a better understanding of natural phenomena by conducting experiments on those phenomena. Unfortunately, in many cases those experiments are too costly and time-consuming and have limited variability.[79]

Instructors can use computers to simulate such science disciplines as thermodynamics. In simulating thermodynamic concepts, they can, for example, compress time intervals to make thermal processes more transparent. Further, they can program the computers to permit students the flexibility to study a number of experimental temperatures and a wide variety of materials. Students can predict thermodynamic occurrences and see if their predictions match the computer simulations, thereby increasing their critical thinking skills and abilities.[80]

Another computer simulation in the science disciplines, "Simulated Qualitative Organic Analysis," focuses student attention on the problem-solving process by simulating a qualitative organic analysis laboratory. Students are tasked to identify "an unknown organic compound using the fewest standard physical and chemical tests and a minimum amount of the compound."[81] Instead of memorizing facts, students use this computer simulation to expand their thought processes and extrapolate on their existing knowledge base.

Science students have yet another software package, "Crystal," to help them understand difficult concepts. Students use this tool to construct, visualize, and analyze three dimensional crystal lattices. They can use these models to calculate such parameters as distances and directions between atoms. Students who have difficulty visualizing three dimensional objects may use "Crystal" to enhance their understanding of molecular principles.[82]

Computers also simulate abstract sociological problems in the real world. The "Social Power" software package allows students to become members of a large multinational corporation. Their goal is to obtain more power within the corporation by managing personal resources and by making deals with other board members. As a result, students analyze power relationships, coalitions, social networks, and strategies to increase their cognition of these abstract skills and abilities.[83]

History courses use computer simulations as well. "The Would-Be Gentleman" simulates social mobility in France during the reign of Louis XIV. Students assume roles within the French bourgeois, acquiring wealth and status by making decisions about investments, personal protectors, suitable marriages, and wills. The students apply their knowledge and understanding of the social history of that time to succeed in the simulation.[84]

History students use the software package "Witches, History, and Microcomputers" to learn how historians interpret the past based on the nature of available data. This computer routine contains information about the chronological, geographic, and socioeconomic dimensions of Salem during the infamous witchcraft trials. Students use a statistical program to investigate relationships and issues the data presents. They develop historical

inquiry skills to interpret the material contained in the software package authentically and make logical conclusions and inferences.[85]

Computer software packages are also used in such humanities subjects as law and English. For example, "Compugraph V. Chang" simulates a legal case in which law students decide how to handle an entire court case from client interviews to courtroom defense. Students use this program to understand legal doctrine, litigation techniques and strategies, and courtroom etiquette.[86]

As for English, an example of a computer software package is "Dickens Web." This tool contains texts, images, and commentaries on Dickens' England. English literature courses use "Dickens Web" to relate aspects of the literary works of that time with information about the authors, the historical settings, and the students' own interpretations and comparisons. Students learn to appreciate multiple perspectives on literature and to recognize the potential relevance of initially disparate information.[87]

Several humanities and social sciences departments at the Air Force Academy also use computer software packages to enhance their learning environments. The English Department recently began using the "Daedalus" software package in its technical writing course. The course director, Lt Col William Newmiller, implemented "Daedalus" on the Academy's networked classroom system (NCS) to take advantage of the networked communication capability between the students and the instructor. "Daedalus" is a tool instructors can use for various facets of language arts: literacy, the writing process, language study, and literature review and discussion. In using "Daedalus", Lieutenant Colonel Newmiller emphasizes in-class communication on the computer and minimizes any verbal communication to help develop the students' writing skills. Networking among the students democratizes the communication process within the classroom, as there is no limit to the classroom discussions and interactions; students who are less likely to participate in oral discussions are more likely to actualize at a computer keyboard. Further, oral communication permits only one speaker at a time, and much of the content of the discussion is lost at the end of the class. Electronically, everyone can type on a keyboard and participate at once; several subdiscussions, in addition to the main discussion, occur simultaneously by splitting windows on the computer screen. The students and instructors also save transcripts of each discussion for later use. The electronic discussion feature of "Daedalus" encourages social interaction while allowing the students to gain a better understanding of the subject matter they are discussing. The students become actively involved in the learning process; not only do they refine their social interaction skills and abilities, but they develop their critical thinking skills and abilities as they present persuading arguments.[88]

The main impact of "Daedalus", though, is its contribution to the quality of student writing. This program "facilitates the process of writing so that students and teachers alike can interact with the text in useful ways that are difficult with paper and pencil."[89] In this role, "Daedalus" is a computer-assisted intervention program with various functions.

In one routine, the program helps students generate ideas and stimulate thought for completing writing assignments by posing sets of questions relating to the assignment and prompting the students to formulate appropriate responses for various types of audiences. Another routine provides the students with a communication situation and asks the students to identify and analyze certain elements and then type in the appropriate responses. The instructor can modify the prompts to fit particular assignments, encouraging the students to alter their thought processes and responses. "Daedalus" also possesses a routine which allows students to read each other's works and respond to specific questions about them. The students can then incorporate the necessary editorial changes before submitting the completed assignments to the instructor. The "Daedalus" program facilitates student learning by focusing on the course objectives; students then respond to the different questions and develop various strategies on the computer to create more comprehensive responses.[90]

"Daedalus" also makes possible joint effort through shared files and through recorded discussions. In addition, the instructor can thoroughly evaluate each student's thought processes and writing skills and abilities by maintaining electronic file folders of each student. These folders may contain written correspondence, formal submissions, and records of student participation in discussions. In the classroom the instructor's role is that of a facilitator, and "Daedalus" frees the instructor from presenting material and permits him or her the opportunity to observe and coach the students. Unfortunately, "Daedalus" requires a labor-intensive effort from the instructor to prepare each lesson. The instructor must research the appropriate material to include in the lesson and then build and refine the presentation. Normally, preparing a "Daedalus" lesson requires significantly more time and effort than preparing a traditional lesson.[91]

The Air Force Academy's History Department uses the NCS in its course on America's first air war. This course immerses students in the content and context of air power during World War I and increases the depth and breadth of understanding of such influences affecting the air war as technology, personalities, and military doctrine. Instructors use the networked computers to develop high-level understanding of the course behavioral and cognitive objectives through the power teaching tool, a program designed to aid the instructors in a lecture-presentation learning environment and to allow students to "embellish" the course lecture materials so that they can develop deeper understandings of the course content and context.[92]

The first function of the power teaching tool is the "power teacher," which serves as a lecture-presentation system for the instructor to use during lectures. This tool electronically displays on each student's computer screen a version of the instructor's guiding questions, which entice students to retrieve prior knowledge, orient their thought processes, and generate discussion. The questions create cognitive dissidence to encourage open discussions and debates. The power teacher supports teaching by allowing the instructor to invoke, display, and manipulate additional information on the screen through

special commands. The teacher can therefore guide the students through the lecture, selecting supporting documents, photographs, and research findings to engage student thinking during the class.[93]

The second function, the "student navigator," is an extended version of the power teacher; students can access it after class hours. They use the student navigator function to access, review, annotate, and explore the lecture material so they can fully benefit from the instructor's guidance questions. In addition, the students can create their own electronic notebooks to retain the instructor's notes, exchange information, and ask the instructor to clarify material. Instructors can examine the questions to determine difficulties students encounter prior to the next lecture.[94]

A final aspect of the History of Air Power course relies on computer simulations to immerse students in various situations which actually occurred during the air war. Using realistic scenarios, the students conduct exercises to develop battle strategies, objectives, and plans of operation. They enhance their presentation skills by showing the results of these exercises in briefings, presentations, and diaries. To learn about and comprehend actual air battles, the students also participate in a commercial simulation program called "Red Baron." This program permits the students to "fly" World War I aircraft on realistic missions. The students maintain logs of their sorties and compare them to diaries and logbooks of actual World War I aviators. The simulations help develop each student's critical thinking and problem-solving abilities.[95]

Communicating through Computers

It may appear that introducing computers in the learning environment eliminates the need for instructors. On the contrary, computers support the instructors; they do not replace them. For example, computers facilitate and improve instructors' communications with students through electronic mail systems. Instructors can use electronic mail to assign work and notify students of schedule changes, grades, and other administrative information which normally takes up valuable class time. Students can prepare and submit assignments through electronic mail, and instructors can review and provide private comments to the students, also through the computer. The electronic postmark features of the system eliminates doubt about the lateness of assignment submitted. Instructors can easily find out if their students have received and read the electronic mail. Transmitted assignments and administrative information are neat, and typewritten documents and hard copies are readily available.[96]

Electronic mail provides more than communications between the instructor and students. "Computer networking and satellite communications technologies can help promote local and long-distance collaboration and communication among students and teachers and can help them become part of the larger world of scholars and scientists."[97] Students can collaborate on problems and assignments with others who are physically located in distant

classrooms, or they can participate in classes while they are home or at other locations. In addition, students are no longer confined to lessons from one particular instructor; experts from around the world can communicate with any number of students simultaneously, much like telephone conference calls. Instructors can use telecommunication systems to teach their specialized subjects to a large and diverse audience. Further, they can instantaneously share information and ideas with each other. Expert instructors can become mentors, sharing their knowledge, skills and abilities in teaching and in their areas of expertise with other instructors from around the country and the world.[98] Instructors can also use these communication systems to build data bases that provide a wide range of resources and information to other instructors. They can obtain information on innovative teaching styles, lesson plans other instructors use, funding sources, public domain software, and summaries of research efforts.[99]

Instructors can use computers to do more than to interact with their colleagues and their students. For example, they can use computer systems to receive data from meteorological satellites to provide meteorology students with real-time information on weather conditions around the world. Students can view storm developments and observe changing weather patterns. Satellite images from the computer can help students conduct real-time analyses and predictions and then allow them to observe how their predictions correspond to reality.[100]

Multimedia Systems

Isolating videodiscs, computer software packages, computer networking, and satellite communications in these discussions of innovative educational technologies does not imply that the technologies are individualized. Rather, these technologies are "increasingly hybrid in nature, combining the features and capabilities of multiple technologies."[101] These multimedia, or hypermedia, systems provide a wider variety of presentation and management options when compared to individual technologies, thus immersing students into a much richer learning environment. They employ text, video, and audio presentations which instructors can control with microcomputers; software tools permit instructors to create links between words, sounds, and images to optimize their presentations.[102]

Multimedia systems that integrate teaching, management, and educational measurement also may enhance the effectiveness of the instructors themselves. Instructors can immediately assess student performance by using complex techniques to provide in-depth analyses of the students' strengths and weaknesses. These measurement systems can, in addition, help instructors to alter existing lessons or develop future lessons to insure that students successfully meet course objectives. Unfortunately, the capabilities of multimedia systems pose a severe limitation to instructors; that is, they are extremely time-consuming endeavors for the instructors. In many cases, they do not have adequate class time to effectively use the systems. In addition,

while visually appealing, many multimedia presentations may not be well suited to the lessons; they may not match the curriculum well. Finally, many multimedia developers assume that all students are self-directed, subject-matter enthusiasts who require little guidance or motivation; clearly, this may not be the case for many students.[103]

Issues Concerning Educational Technologies

The capabilities of innovative educational technologies in the learning environment make it difficult to imagine why anyone would refuse to introduce these technologies into the classroom. However, four issues address the resistance of many educators in using innovative educational technologies in the classroom: the cost of educational technology; the time educational technology requires of students and teachers; the influence of different presentation modes on student learning; and the instructor's role in the classroom.

A universal concern of educators engaged in enhancing teaching styles is the cost of innovative educational technologies. Obviously, purchasing and installing the hardware and software required of such technologies as videodiscs, computers, and networked computer systems is quite expensive. However, these expenses are "frequently only half of the total initial cost to establish a facility; other costs include renovation of the room and utilities, and software and supplies. In addition, these initial costs are compounded by annual expenses for software, supplies, utilities, equipment maintenance, and user support services."[104] Further, instructors must be trained to use these technologies; training can accrue from workshops or from individual training sessions. Educators and administrators must therefore consider the costs of training the instructors as part of the cost of the educational technologies.

In some cases, however, innovative educational technologies may reduce the cost of teaching. For example, educators can use media repeatedly to demonstrate expensive or dangerous experiments; in this case, they would actually perform the experiment once and record it for future use. Still, before institutions commit to implementing innovative educational technologies in their courses, educators and administrators must consider if the benefits of these technologies are worth the costs they incur.

Another issue to consider is the time required for both teachers and students to master innovative educational technologies. On one hand, educational technologies increase the effectiveness of teachers and students. For example, if using computers as tutors reduces the amount of time students need for learning and decreases the time instructors require to tutor students, the resulting free time subsequently reduces the workload of both instructors and students.[105] Further, when students become immersed in the innovative educational technologies during the lesson, the instructors can devote more time to monitoring individual students and providing instant extra tutoring when needed.[106] Electronic mail also functions as an effective time management tool by helping students to avoid standing in long lines outside an instructor's door solely to ask for clarification of a problem or to learn

the results of a test. Instructors don't have to contend with students' insistence on individualized attention at inconvenient times, and they don't have to spend time arguing with students over misplaced or late submissions.[107]

Educational technologies, however, sometimes requires additional time from instructors and students. For example, the computer networks they use for electronic mail messages can experience system failures or suffer from software and hardware inefficiencies, which can make them time-consuming for the instructors and students.[108] More importantly, though, as Karen Sheingold states, "Once teachers begin to use technologies well to advance student learning, they often (1) need more time to learn about, obtain additional training in, and plan for the use of the technologies; and (2) want students to have longer blocks of time in which to do their technology-based work."[109] To incorporate innovative educational technologies in a course, instructors must devote a tremendous amount of time to deal with such technical requirements as hardware and software capabilities and limitations, room design, and efficient utilization of the technologies. They also must give a great deal of attention to designing and implementing student activities which use the technologies. Even after these technologies are in place, they continue to demand instructors' time to incorporate changes; the instructors must research and implement these changes.[110] The time involved in learning to use innovative educational technologies also can affect students adversely. When they are introduced to a new media as a teaching tool, students must confront the burden of having to learn and to adapt to this new technology. Hopefully, however, these requirements have only short-term impacts on each student's time; as the students gain familiarity with a particular system, they spend less time learning it and are better able to use its time-saving shortcuts.[111]

If educators and administrators accept the costs associated with innovative educational technologies and compensate students and instructors for the additional time these educational technologies require of them, they may notice that the manner in which the technologies present material enhances the learning process. The most dramatic difference between traditional teaching methods and teaching incorporating innovative educational technologies is the variety of ways instructors can present lessons and activities. Simon Hooper and Michael Hannafin believe that "the modes of presentation may include the individual or combined presentation of sound, still or motion picture, text, graphics, animation, and computer-generated sound."[112] Different methods of presenting material may enhance student learning and comprehension by reducing boredom and increasing interest and, consequently, involvement.

There are, however, disadvantages to the presentation modes innovative educational technologies introduce. Studies have shown that "text is read more slowly . . . and that comprehension is lower . . . when read from the computer screen than from print-based media."[113] Apparently, using the computer screen as a presentation mode may inadvertently detract from the intended learning process.

Further, the type of presentation either focuses or diverts attention from the learning task; poorly designed or inappropriate presentations distract the students from the intended learning. The manner in which material is presented may affect the level of student arousal. The presentations also affect how intensely students process the material presented in the lesson.

For example, some students perceive video images as too easy and won't devote much mental effort to the lesson; movies and television are passive presentation modes, and students may not become involved in the lesson.[114] In addition, some students may become intent on studying computer animations and fail to concentrate on the lesson. As for videodisc scenarios, students sometimes watch the scenery and people in the video rather than paying attention to the lesson. Hence, not all presentation modes equally motivate students; instructors must consider the intended audience. Reading printed material requires more mental effort than television, but this effort increases the student's ability to make inferences. Most presentation modes enhance student learning under specific conditions: the mode must be congruent with the learning task; students must be cued to attend to the presentation; and the content difficulty of the presentation must be optimized.[115]

Still, there will be students who are indifferent to the mode of presentation; while many students prefer the presentations offered by innovative educational technologies, others prefer the traditional teaching methods. Situations exist where the learning environment or the presentation mode does not influence student learning. Rather, a student's ability to learn may be influenced by such motivational factors as interest in the topic and the type of course the student is enrolled in; that is, a required or an elective course.

More than just the topic or the type of course affects student motivation. Indeed, the instructor has a great influence on the attitudes and motivation levels of students. Teachers who are well trained in using innovative educational technologies and have positive attitudes about these technologies can help students to improve their attitudes. They devote the necessary time and energy to use these innovative educational technologies effectively to insure that students benefit from the lessons.[116]

Conversely, teachers who are poorly trained or those who harbor prejudices against using educational technologies in a classroom environment "may communicate negative attitudes to their students (which may lessen motivation); such teachers may also become defensive or derisive about student's interest in using" the technologies.[117] Critics claim that educational technologies "displace teacher's professional skills, deprive teachers of their opportunity to 'perform', and reduce their participation in the student's learning."[118] Hence, when compelled to use innovative educational technologies in the classroom, some instructors tend to use the technologies to teach in the same manner as they are accustomed to; the technologies simply make their current teaching styles more efficient. Learning still focuses on the instructor rather than the student, and instructors continue using innovative educational technologies improperly in a traditional teaching environment.[119]

However, the critics of innovative educational technologies overstate their case when they contend that it reduces the need for skilled teachers. As teaching becomes more technical, instructors must adapt to a changing learning environment to continue helping students learn. Using technology in the classroom increases the complexity of teachers' jobs, as they have available more complex approaches to teaching.[120] Teachers become "facilitators who help students construct their own understandings and capabilities in carrying out challenging tasks. This view puts the emphasis on the activity of the student rather than on that of the teacher."[121] The role of the teacher is more challenging: instead of dictating to students the answers to posed problems, the teacher "requires more subject-matter expertise and more skill in guiding students to derive appropriate conclusions from an activity."[122] In addition, instructors must be skilled in motivating students to use the innovative educational technologies, especially those students who possess negative attitudes towards the technologies.

Courses using innovative educational technologies thus require skilled teachers. Of course, teachers using traditional teaching methods also must be skilled. The effectiveness of a particular learning environment is therefore teacher-dependent. Similarly, the appropriateness of the learning environment depends on both the type of course the instructor teaches and the individual student enrolled in the course. Some courses and students are better suited for traditional teaching methods; others are better suited for teaching styles which use innovative educational technologies.

Summary

This chapter presents an unbiased view of two categories of teaching styles: one which is based on traditional teaching methods and the other which uses innovative educational technologies as aids to teaching. There are numerous advantages and disadvantages to either category of teaching style; I do not wish to judge which teaching style is more effective in teaching students. Rather, I point out that the methodology or methodologies educators use to assess and compare the performance of students in different learning environments must be robust enough to account for the complexities of these two categories of teaching styles. Otherwise, the methodologies will undoubtedly bias the results of the comparisons. The next chapter examines existing studies which compare various styles of teaching.

Notes

1. *Webster's II New Riverside University Dictionary* (Boston: The Riverside Publishing Company, 1988), 1187.

2. Dr Michael J. Hannafin, "Emerging Technologies, ISD, and Learning Environments: Critical Perspectives," *Educational Technology Research and Development* 40, no. 1 (1992): 57.

3. Ibid., 58.
4. Ibid.
5. Ibid., 59.
6. Ibid.
7. Ibid.
8. Ibid., 60.
9. Bruce Joyce and Marsha Weil, *Models of Teaching* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972), 8.
10. Ibid., 9.
11. Ibid.
12. Ibid., 10.
13. Ibid.
14. Frederick Hill, "The Cult of Newness," *American School and University* 61, no. 10 (June 1989): 13.
15. Ibid.
16. Gordon Kirk, "Introduction: Techniques of Teaching and Assessment," in *Handbook of Educational Ideas and Practices,* ed. Noel Entwistle (London: Routledge, 1989), 783.
17. Helen Cowie and Jean Rudduck, "Learning Through Discussion," in *Handbook of Educational Ideas and Practices,* 804–5.
18. Jim W. Corder, "Traditional Lectures Still Have a Place in The Classroom," *The Chronicle of Higher Education* 37, no. 39 (12 June 1991): 32.
19. Kirk, 783.
20. Barbara Kane, "Undergraduate Students' Concerns and Preferences Regarding Structured and Unstructured Lecture-Discussions," *Education* 111, no. 1 (Fall 1990): 117–24.
21. E. Peter Volpe, "The Shame of Science Education," *American Zoologist* 24 (1984): 433.
22. Corder, 32.
23. Ibid.
24. Kane, 117–24.
25. Ibid., 117.
26. Ibid.
27. John Biggs, "Teaching for Desired Learning Outcomes," in *Handbook of Educational Ideas and Practices*, 689.
28. Ibid.
29. Marion Blythman and Donnie Macleod, "Co-Operative Teaching," in *Handbook of Educational Ideas and Practices*, 853–55.
30. Hannafin, 54.
31. Ibid.
32. Cowie and Rudduck, 803.
33. Patricia Kennedy Arlin, "Teaching as Conversation," *Educational Leadership* 48, no. 2 (October 1990): 82–84.
34. Cowie and Rudduck, 804.
35. Ibid., 806–7.
36. Leonard P. Oliver, "Study Circles: Individual Growth Through Collaborative Learning," *New Directions for Adult and Continuing Education*, no. 53, Spring 1992, 85–87.
37. Ibid., 87–95.
38. Cowie and Rudduck, 808–9.
39. Morry Van Ments, "Simulations, Games and Role-Play," in *Handbook of Educational Ideas and Practices*, 825.
40. Ibid., 823–24.
41. Ibid., 827.
42. Ibid., 827–31.
43. Gerald Collier, "Syndicate Methods," in *Handbook of Educational Ideas and Practices*, 813–15.
44. Frank B. Murray, "Co-operative Learning," in *Handbook of Educational Ideas and Practices*, 859.
45. Ibid., 860–62.

46. Collier, 817–19.
47. Volpe, 437.
48. Ibid.
49. Ibid., 437–38.
50. Gaye Manwaring, "Open Learning," in *Handbook of Educational Ideas and Practices*, 842–44.
51. Ibid., 843–47.
52. Ibid., 848–49.
53. Corder, 32.
54. Andrew Trotter, "Computer Learning with Integrated Systems," *The Education Digest* 56, no. 4 (December 1990): 28.
55. Ibid.
56. Dr Michael Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), interview with author, 12 August 1992.
57. Ken McCullough, "Technology, Decades Old and Still Growing!" *Adult Learning* 4, no. 5 (January/February 1993): 5.
58. Ibid.; Lynn Olson, "Use of Technology in Schools: Still Elusive," *The Education Digest* 57, no. 9 (May 1992): 21.
59. Robert B. Kozma and Jerome Johnston, "The Technological Revolution Comes to the Classroom," *Change* 23, no. 1 (January/February 1991): 10.
60. James E. Dezell, Jr., "Technology in the Classroom," American School and University 63, no. 6 (February 1991): 21–24.
61. Hannafin interview.
62. Dezell, 21–24.
63. Raymond Kurzweil, "Learning in the Age of Knowledge," *Library* 116, no. 19 (15 November 1991): 62.
64. Karen Sheingold, "Restructuring for Learning with Technology: The Potential for Synergy," *Phi Delta Kappan* 73, no. 1 (September 1991): 20.
65. Kozma and Johnston, 16.
66. Ibid.
67. Ibid., 18.
68. Ibid.
69. Maj Judith Brisbois, Department of Foreign Languages, Air Force Academy (HQ USAFA/DFF), interview with author, 12 August 1992.
70. Amy Rose, "Technology—Servant or Master?," *Adult Learning* 4, no. 3 (January/February 1993): 6.
71. Major Brisbois interview.
72. Allan Collins, "The Role of Computer Technology in Restructuring Schools," *Phi Delta Kappan* 73, no. 1 (September 1991): 29.
73. Trotter, 28.
74. James A. Kulik and Robert Bangert-Drowns, "Computer-Assisted Learning," in *Handbook of Educational Ideas and Practices*, 834.
75. Ibid., 834–35.
76. Rose, 6.
77. Ibid.
78. Kozma and Johnston, 15–16.
79. Eileen L. Lewis, Judith L. Stern, and Marcia C. Linn, "The Effect of Computer Simulation on Introductory Thermodynamics Understanding," *Educational Technology* 23, no. 1 (January 1993): 45–46.
80. Ibid.
81. Kozma and Johnston, 18.
82. Ibid., 17.
83. Ibid., 16.
84. Ibid.
85. Ibid., 18.
86. Ibid., 16.

87. Ibid., 17–18.

88. Lt Col William Newmiller, Department of English, Air Force Academy (HQ USAFA/DFENG), interview with author, 11 August 1992.

89. Kozma and Johnston, 14.

90. Newmiller interview.

91. Ibid.

92. Dr Kathleen Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), and Capt Michael Grumelli, Department of History (HQ USAFA/DFH), interviews with author, 13 August 1992.

93. Ibid.; "Power teaching tool" notes from Dr Kathleen Hannafin, 13 August 1992.

94. "Power teaching tool."

95. Grumelli interview; "Draft History 495 Course Syllabus," 13 August 1992.

96. Heinz V. Dreher, "Electronic Mail: An Exemplar of Computer Use in Education," *Educational Technology* 24, no. 8 (August 1984): 136–37.

97. Sheingold, 20.

98. "Long-Distance Teaching," *The Futurist* 24, no. 6 (November/December 1990): 48.

99. John Fleischman and Gerald H. Kilbert, "Adult Education Technology in the Golden State," *Adult Learning* 4, no. 3 (January/February 1993): 15.

100. Jamie B. Spigler, "Computers Enhance Classroom Teaching," *American School and University* 64, no. 6 (February 1992): 31–32; Dr Kathleen Hannafin interview.

101. Michael J. Hannafin, Kathleen McDermott Hannafin, and David W. Dalton, "Feedback and Emerging Instructional Technologies," draft chapter for *Feedback and Interactive Instruction*, ed. John V. Dempsey and Gregory W. Sales (Englewood Cliffs, N.J.: Educational Technology Publications, 1991), 4.

102. C. V. Bunderson, J. B. Olsen, and A. Greenburg, *Computers in Educational Assessment: An Opportunity to Restructure Educational Practices* (Pennington, N.J.: The Institute for Computer Uses in Education, 1990), iii; Ward M. Cates, "Instructional Technology," *The Clearing House* 65, no. 2 (November/December 1991): 70.

103. Cates, 70.

104. Kozma and Johnston, 22.

105. Kulik and Bangert-Drowns, 839.

106. Spigler, 31–32.

107. Dreher, 38.

108. Lt Col John Souders, director of research, Directorate of Education, Air Force Academy (HQ USAFA/DFER), interview with author, 20 January 1993.

109. Sheingold, 21.

110. Kozma and Johnston, 19–20; Olson, 21.

111. Dr Kathleen Hannafin interview.

112. Simon Hooper and Michael J. Hannafin, "Psychological Perspectives on Emerging Instructional Technologies: A Critical Analysis," *Educational Psychologist* 26, no. 1 (1991): 80.

113. Michael J. Hannafin and Simon Hooper, "An Integrated Framework for CBI Screen Design and Layout," *Computers in Human Behavior* 5 (1989): 155.

114. Souders interview.

115. Hooper and Hannafin, "Psychological Perspectives," 78–80.

116. Mark Sherry, "Implementing an Integrated Instructional System: Critical Issues," *Phi Delta Kappan* 72, no. 2 (October 1990): 118–19.

117. Ibid., 119.

118. Trotter, 30.

119. Jane L. David, "Restructuring and Technology: Partners in Change," *Phi Delta Kappan* 73, no. 1 (September 1991): 80; Dr Michael Hannafin interview.

120. Trotter, 32; David, 79.

121. Collins, 29.

122. Kozma and Johnston, 19.

Chapter 3

Research Comparing Styles of Teaching

The previous chapter briefly discussed various traditional teaching methods and different ways in which instructors use innovative educational technologies in the learning environment. Before developing methodologies to compare the effectiveness of different teaching styles, it seems logical to first review existing comparison studies. Hence, this chapter briefly summarizes a representative sampling of studies educators have already conducted to compare various styles of teaching.

Most of the existing studies have focused on comparing traditional teaching methods to teaching using educational technologies. Few researchers have conducted extensive studies of, for example, the effectiveness of didactic teaching styles compared to such novel traditional teaching methods as seminars or role-playing. In addition, few studies compare the effectiveness of one innovative educational technology to another. Consequently, a plethora of research topics lies untapped; researchers can contribute tremendously to the field of education by conducting studies which compare the effectiveness of different teaching styles regardless of their classification as either traditional teaching methods or teaching styles which use innovative educational technologies as teaching aids. Unfortunately, such studies are scarce. Therefore, the remainder of this chapter summarizes those studies that compare traditional teaching methods and teaching using educational technologies.

Categories of Studies Comparing Traditional Teaching Methods and Teaching Using Educational Technologies

The studies which compare traditional teaching methods and teaching using educational technologies can be divided into three general categories. The studies in one category obtain information about the effectiveness of a particular innovative educational technology by comparing different teaching media. These studies determine which technology, medium, procedure, or teaching style best serves a given course or set of lessons. A second category of studies increases the understanding of how different innovative educational technologies function and what psychological effects they have on students. These studies focus on how technologies function psychologically rather than on how their effectiveness in aiding teaching compare to other media. Studies

in the third category enhance the educational process by identifying ways to evaluate and improve technologies. Rather than compare a certain technology to another or to a different teaching style, these studies demonstrate how to improve existing technologies.[1]

While these three categories of studies help educators understand the role of innovative educational technologies in the field of education, I examine only those studies that fall within the first category mentioned above. I highlight a few of the many studies conducted to reflect the quality of the research educators have accomplished.

Studies Comparing Media

The first generation of research studies on using media as educational technologies in the classroom occurred in the 1920s; hence, the data base of research examining educational technology spans over seven decades. As early as 1924, researchers conducted a systematic series of experiments with film and concluded that the effectiveness of verbal teaching, when compared to teaching using visual media, depended on the nature of the style of teaching and on the learner's previous experience. However, researchers did not pursue this idea of an interaction between media, learning processes, and students; instead, later studies usually focused on finding "the best medium."[2]

For example, a 1968 study examined the effects of motion in film on students' learning. According to Gavriel Salomon and Richard Clark,

> Three types of learning tasks were studied, and 582 students, differing in age, sex, ability, and specific knowledge, were tested. The researchers reached the conclusion that motion in films facilitated learning more than still pictures. . . . This was the case regardless of learner or task differences.[3]

While the differences between how this study and the 1924 study were conducted are relatively minor, the conclusions are diametrically opposed. Specifically, the 1924 study compared verbal teaching to teaching using visual media and concluded that effectiveness depended on the teaching styles and the students rather than the media itself. Conversely, the latter study compared two different types of visual media and concluded that a distinct difference existed between the effectiveness of the two media. An anomaly clearly exists in the conduct of both studies.

A study conducted in 1972 also focused on finding the best medium; in this case, researchers compared the relative effectiveness of radiovision and television in a mathematics course. The qualities of radiovision, which is a tape-recorded sound track accompanied by a filmstrip, permitted students to employ repetitions, discussions, and other time-intensive teaching aids. While the television presentation contained extensive animations, it did not allow for interruptions. In this experiment, researchers found no significant differences between the two groups; researchers concluded that the less costly radiovision medium facilitated learning as much as television did.[4] These

conclusions contradict those from the 1968 study but support the 1924 study. Salomon and Clark examined the 1972 study's internal validity, which is the extent to which researchers focus the study by eliminating such external influences as environmental variables and student and instructor characteristics. They found that the study sacrificed internal validity by failing to identify and eliminate "a host of known and unknown variables [that influenced the study]. Even if results were to favor one medium, no clear explanation could be provided, nor would it be possible to suggest which of all the participating variables is responsible for the outcomes."[5]

Studies Comparing Media Attributes

One research approach addresses increasing the internal validity of such studies by comparing various media attributes. These attributes embrace the "capabilities of each medium and are properties of stimulus materials which are manifest in the physical parameters of media."[6] Media researchers specify relevant attributes of the different media and compare these attributes to determine which medium of teaching is more effective.

As an example, a media researcher might want to find out if a motion picture is more effective than a textbook. The researcher describes the motion picture and textbook in terms of their attributes and conducts a study comparing these attributes to determine which media is more effective. Unfortunately, specifying the appropriate attributes can be difficult. The term *motion picture* itself is inexact; it can encompass such diverse attributes as silent or sound, color or black and white, animated or live action, and high-speed or time-lapse photography. The researcher must determine which of these attributes most directly influences student performance in the classroom. Further, many of the motion picture attributes do not relate to the attributes of a textbook, which impedes comparing the two media. Hence, the researcher must focus on establishing specific attributes which not only influence the learning process but describe the different media he or she is examining.[7]

For example, a 1973 research study focusing on comparing media attributes examined the relative merits of words and pictures in simple learning tasks. Researchers discovered that people have an astonishing recognition memory for pictures, and, consequently, found that pictures were superior to words for eliciting recall and recognition of items in paired-associate learning. Pictures also surpassed words for helping students to learn concrete subject matter and for providing to students who lacked verbal equivalents for the pictures a means to comprehend the subject matter. Words, on the other hand, out performed pictures in terms of a student's comprehension of abstract subject matter and channeled his or her thought processes more narrowly and predictably. While instructors could use pictures as prompts so students could recognize unfamiliar words in textual material,

researchers found that pictures occasionally miscued the students or diverted their attention from textual material. In fact, adding pictorial embellishments to verbal material increased student enjoyment and interest but did not necessarily increase the students' comprehension of the topic addressed in the text.[8]

A similar study compared auditory and visual presentations of textual material and concluded that reading was superior to listening. The results of college students taking a multiple-choice retention test showed that reading was superior to lecture, television lecture, and radio. Reading also surpassed listening when college students studied moderately and highly difficult material; this conclusion was especially true when students took advantage of the preferability of print. In addition, auditory presentations exceeded visual presentations for short-term learning tasks at high-presentation rates. Visual information recall improved with slower presentation rates. Researchers also determined that simultaneous audiovisual presentations of words were not superior to learning from print alone, but students learned more when instructors presented pictures in conjunction with the audiovisual presentation. As conclusive as the researchers were in this study, they also stated that several studies showed no difference between auditory and visual presentations, while other studies demonstrated that auditory presentations were superior to visual presentations, and vice versa.[9]

Studies of Television and Videodisc Instruction

Studies of the effectiveness of television as a visual medium for education have produced results as diverse as those summarized above. Most researchers have found no difference between students exposed to teaching styles using television and those taught in the traditional manner. However, the conclusions of a 1962 study suggested that teaching using television was more effective than face-to-face (or lecture) teaching in mathematics and the sciences and less effective in the humanities. Further, a later study proved videotapes to be more effective than face-to-face teaching when instructors used them to teach such performance skills as typing, sewing, and athletics.[10] Studies conducted in the early 1980s comparing television presentations to those using print showed that television required less effort but that print improved the student's ability to make inferences. Implied, rather than supplied, images increased a student's mental effort and led to "more elaborations and deeper processing."[11]

Technology has improved the capabilities of the television medium through the introduction of interactive video systems into the classroom. Naturally, researchers have conducted studies to assess the effectiveness of this educational technology. A 1989 study compared an interactive video system to motion picture films for use as teaching aids in an eighth grade science class. While both educational packages contained the same material, students

exposed to the interactive video system controlled the pacing of the presentation. However, rather than having each student interact with an individualized video system, the entire class used a single system. Hence, the students and the instructor interacted to maintain a presentation pacing that satisfied everyone. The instructor administered pre- and posttests both to the students using the video system and the students exposed to motion picture films. Researchers used the pretest scores, as well as such student characteristics as sex, race, and socioeconomic status, to match students between the groups, and they computed the measure of each medium's effectiveness by subtracting pretest scores from posttest scores and by comparing the results from each matched student-pair. The results showed that the students exposed to the video system had significantly higher test score gains than the students exposed to the motion pictures. Based on a cost-effectiveness evaluation, interactive video also was significantly more efficient than the motion picture films as a teaching aid.[12]

A 1992 study altered the previous study's focus by investigating "the relationship between students' perceptions of the effectiveness of interactive television as a medium of [teaching] and their end-of-course evaluations of two instructors who prefer different teaching styles."[13] The students tested were enrolled in a semester-long interdisciplinary undergraduate course co-taught by two instructors who employed an interactive television (ITV) system. The instructors broadcast their classroom presentations live from an on-campus broadcast classroom to three off-campus classrooms. Students completed an end-of-course evaluation while the instructors completed a teaching style inventory.[14]

The instructors' inventories indicated that one instructor preferred an instructor-oriented teaching style while the other preferred a student-oriented style. The students' evaluations indicated that students tended to rate ITV systems and instructors separately; students separated their perceptions of the teaching medium from their perceptions of the instructor's effectiveness. About one-half of the students believed that the ITV system was an effective way to teach the course, and they had little preference in the mode of teaching (either on-site or at an off-site classroom). Further, the students in the study rated both instructors as effective but preferred the instructor-oriented style over the student-oriented teaching style. The study results showed that the students' evaluations of instructor effectiveness were not adversely effected by their satisfaction levels with an ITV system.[15]

Studies of Computer-Based Systems

Research has not been confined solely to assessing the effectiveness of visual media; researchers have examined other educational technologies. Computer-assisted instruction, a drill-and-practice tutor for students, is one

such technology. Numerous studies have examined various aspects of computer-assisted instruction.

For example, one study conducted between 1976 and 1986 focused on determining the effectiveness of computer-based tutorials. Students enrolled in an undergraduate introductory biology course used the tutorials to supplement lectures and regular assignments. The tutorial's multiple-choice format simulated typical exams and required students to synthesize knowledge extracted from the text and lectures to solve problems. The tutorial provided explanations for right and wrong answers, helping students to modify their scientific reasoning abilities. The study showed significant improvements in student scores on the final examinations in the 10 years since instructors introduced the tutorial. While the final examination varied each year, it always covered the same principles, implying a common means of comparison for 10 years. The students used the tutorials voluntarily, and the study revealed that the amount of time the academically weak and strong students devoted to the tutorial was directly related to their performance on the final examination. Finally, the study noted that students were more enthusiastic about using the computer-based tutorial as opposed to attending review sessions or receiving packages of practice exams with rationales for the answers typed beside the questions.[16]

Mastery learning, a variation of computer-based tutorials, teaches ordered skills by systematically teaching, testing, and employing remedial instruction until students achieve established performance levels. In 1988 David Dalton and Michael Hannafin examined the effects of mastery learning by using combinations of teacher- and computer-based initial and remedial teaching techniques. Teachers initially taught students in an eighth grade mathematics course an algebraic computation lesson and then administered a mastery quiz; those students who did not pass the quiz received a remedial lesson the following day. Researchers divided the students into four groups: the instructors taught both the initial and remedial lessons to the first group; the instructors taught the students in the second group the initial lesson and the students used remedial drill-and-practice, computer-assisted instruction for the remedial lesson; students in the third group used the computer for both their initial and their remedial lessons; and students in the final group used the computer for their initial lesson, while the instructors taught the remedial lesson. The researchers compared the students' quiz results to the results of quizzes administered to a control group of students who did not participate in the mastery learning program. Although the students in the mastery learning program performed significantly better than the students in the control group, neither the computer-assisted instruction nor the teachers' teaching influenced student performance significantly during the initial or remedial lessons. However, the study concluded that varying the method of delivery, whether from teaching by the instructors to remedial instruction on the computer or vice versa, provided the greatest impact on quiz performance.[17]

Computer-assisted instruction also has been used to help high school students prepare for the Scholastic Aptitude Test (SAT). Researchers

conducted a study in 1990 to determine if computer-assisted instruction, drill, and practice were viable techniques to improve SAT scores when compared to traditional coaching techniques. Eight students volunteered to use the computer to prepare for the SAT while eight other randomly selected students served as a control group and used traditional methods to prepare for the test. The study results indicated that students attained higher SAT scores by using computer-assisted instruction, implying that it helped to prepare students for the SAT more effectively than other methods.[18] While these results are interesting, they may not have statistical significance due to the small sample size of students the researchers used in the study.

Researchers also have compared courses which use computer simulations with traditional courses. One study, for example, investigated students' achievements in chemistry and their attitudes towards the chemistry discipline that resulted from two instructors' simulations of chemistry experiments on a computer. Researchers randomly assigned 200 high school students to six classes; the two chemistry instructors used different styles of teaching in each of three classes. They used the same lecture and guided discussions to teach all the students; hence, the differences in the classes lay in the computer-simulated and laboratory experiments they taught.[19]

Two control classes participated in six laboratory activities in which they examined the mole concept, chemical reactions, gases, and solutions. The students followed detailed instructions to conduct specific experiments and to collect the required data. They then compared their experimental findings with approved solutions. The instructors exposed two classes to a problem-solving approach to experimentation, and the students used this approach to examine the same laboratory activities as the control classes. The instructors gave the students laboratory sheets which either provided a series of possible solutions to each activity or left the construction of the solution to the students. The students designed and conducted their own experiments to investigate the activity, interpreted their collected data, and drew conclusions based on the experiment's results and hypotheses they had formulated. The final two groups also participated in the same laboratory activities but used computer simulations to conduct the experiments. Their approach resembled that of the problem-solving group, but this group of students could alter certain aspects of their experiments and repeat them if desired.[20]

To compare the performance of these students, researchers administered pretests to determine the students' prior knowledge of chemistry, attitudes towards the chemistry discipline, and scientific inquiry skills. They found no statistically significant differences between any of the three groups of students prior to the study. The researchers administered posttests and examined the students' growth and achievement resulting from their enrollment in the course. They found no differences between the students in the problem-solving and computer simulation classes with respect to their knowledge and comprehension of specific chemistry principles. However, their performances were significantly greater than the control classes. The researchers found these same results when they examined the students'

scientific inquiry skills. Finally, they determined that the students in the computer simulation classes had significantly more positive attitudes towards chemistry than the other two groups, and the control group had the more negative attitudes.[21]

Based on these results, the researchers concluded that the computer-simulated experiment and problem-solving approaches to chemistry produced significantly greater achievement and scientific inquiry skills than the traditional approach to teaching chemistry. They determined that computer simulations were more effective than traditional laboratory experiments, because students could control the laboratory conditions and did not have uncontrolled variables or measurement errors. Further, students may have understood concepts better, since they could reexamine segments of the experiments, and they had a variety of situations which could alter the experiments and permit them to test their hypotheses. Finally, the researchers concluded that students may have perceived the computer-simulated experiments as games which may have engendered more positive attitudes towards chemistry.[22]

The researchers' conclusions need to be interpreted with the caveat that they were dealing with a small sample size in their study; as was the case in previous studies, the results of the study may not be statistically significant. Further, the researchers cannot infer the above conclusions when, in the study, they found no statistically significant difference between the achievements and the scientific inquiry skills of the computer-simulation and problem-solving groups.

Synthesizing the Results of Multiple Studies

The above studies suggest that researchers arrive at quite diverse conclusions when they employ a variety of approaches to assess the effectiveness of teaching styles and innovative educational technologies. In addition to the different media applications they investigate, studies differ in experimental designs and settings. Consequently, the results frequently differ. Researchers often synthesize the results of multiple studies to find common threads and plausible explanations for the different conclusions.[23] Two basic methods of data synthesis exist: box-score reviews and meta-analyses. According to James Kulik, Robert Bangert-Drowns, and George Williams,

> Box-score reviews usually reported the proportion of studies favorable and unfavorable toward computer-based instruction, and often provided narrative comments about the studies as well. . . . Meta-analysts used (a) objective procedures to locate studies; (b) quantitative or quasi-quantitative techniques to describe study features and outcomes; and (c) statistical methods to summarize overall findings and to explore relationships between study features and outcomes.[24]

Box-score reviews limit the amount of information a researcher can convey, which is a critical disadvantage in the research on the effectiveness of

teaching styles. Conversely, meta-analysis statistically analyzes a large collection of research results from various studies and integrates their findings. Researchers must characterize the features and outcomes of the selected studies by using multivariate techniques to relate the characteristics of the studies to the outcomes. Researchers usually express the results in terms of the difference between the means of two groups of people divided by a standard deviation common to both groups.[25]

Box Score Reviews

An example of box-score reviews demonstrated that the results from 10 independent studies showed a substantial advantage to using computer-assisted instruction. Other studies concluded that computer-assisted instruction, when complementing regular teaching, either improved achievement test scores or were as effective as traditional teaching methods. While computer-assisted instruction often produced better results than traditional teaching methods on final examinations, they produced worse results on retention exams. Still, computer-assisted instruction reduced the amount of study time required by students.[26] While these results are interesting, they express generalizations and educators can therefore subject them to different interpretations. Further, box-score reviews do not consider differences in the studies' designs or media applications, and without a common baseline, researchers cannot realistically compare the results.

Meta-Analysis

Meta-analysis is an alternative to synthesizing research results in the form of box-score reviews. Probably the first study to apply the meta-analysis technique was accomplished in 1977; it focused on elementary and secondary school students. Elementary school students using computer-based instruction in mathematics courses raised their examination scores by 0.41 standard deviations, while high school students raised their mathematic examination scores by 0.3 standard deviations. At the elementary level, this increase equated to raising the examination scores from the 50th percentile to the 66th percentile. The study reported that while these results were not as large as those obtained when students used peer and cross-age tutoring programs, the results were significantly greater than those produced when the students used programmed instruction or individualized learning packages.

Later meta-analysis syntheses revealed that college students raised their examination scores by between 0.1 and 0.25 standard deviations when they used computer-assisted instruction, or roughly an increase from the 50th percentile to between the 54th and 59th percentiles in achievement. Computer-assisted instruction also substantially reduced the amount of study time students required. These results suggest that, at least in mathematics education, the effectiveness of computer-assisted instruction results from a function of the student's learning level. At the introductory levels of learning, the computer provides an interactive medium to stimulate and guide

students, but at higher learning levels, this interaction may be unnecessary or even obstructive.[27]

A 1982 meta-analysis research effort focused on 51 studies that examined computer-assisted instruction in grades 6-12. Variables and categories described essential features of the studies. Several variables delineated such computer applications as drill and practice, tutoring, computer-assisted teaching, and simulations. One variable indicated whether the computer complemented or replaced traditional teaching, while another indicated the study duration. Other variables defined such aspects of the experimental design of the studies as assigning students to experimental and control groups randomly or nonrandomly, using standardized examinations, and having a single teacher instruct both the experimental and control groups. Other variables described miscellaneous features of the course settings, including student ability levels and subject matter. Eventually, researchers used 12 variables to explain variations in the outcomes of the studies.[28]

The meta-analysis summarized findings in three areas: student performance on final and retention examinations; attitudes towards the subject matter, computer, and instruction; and study time. A term called effect size, which was "the difference between the means of two groups divided by the standard deviation of the control group," quantified the outcomes of these areas.[29] The meta-analysis showed that computer-assisted teaching raised final examination scores by .32 standard deviations, from the 50th percentile to the 63d percentile; it also increased the scores on retention examinations. Students had positive attitudes towards computer-assisted teaching and the computer-based courses they took. Finally, using computers reduced the amount of time students needed to learn the course material.

Two conclusions concerning the 51 studies themselves are of particular interest. First, researchers reported that the more recent studies indicated that computer-assisted teaching had stronger effects on student performance; it appears that instructors have more appropriately used innovative educational technologies recently. Second, researchers concluded that studies of short duration produced stronger effects on student achievement when compared to studies of longer duration; shorter studies may be better controlled and, consequently, more likely to estimate true effects.[30] Conversely, studies of longer duration may capture long-term effects of the learning process on student achievement. Shorter duration studies may produce anomalous effects which researchers could average out if the studies were of longer duration.

Another study using meta-analytic procedures to analyze computer-assisted teaching in college environments reported an increase of about 0.5 standard deviations on final examination performance when compared to traditional teaching. This result translates into an increase from the 50th to the 66th percentile in examination scores. However, when one instructor planned and presented both the computer-based and traditional courses, researchers noted that the effect size dropped to 0.13 standard deviations. The decrease in effect size implies that the larger effect size resulted from

such uncontrolled variables as course content, novelty, and styles of teaching rather than from the medium of computer-assisted teaching.[31]

Researchers also used meta-analysis to integrate the findings of 74 studies comparing visual-based teaching to traditional teaching in colleges. The studies examined seven visual-based teaching applications: still projection; film; multimedia; closed-circuit television; educational television; video used for observation; and video used for feedback. They focused on various courses, ranging from the sciences to the humanities. Research designs of individual studies also differed; some were loosely controlled while others were rigorous. The researchers defined 16 variables, and study outcomes were quantified as one of five types: examination scores, retention scores, correlation between aptitude and achievement, course ratings, and course completion. A variation of the previously defined effect size quantified the outcomes; for this analysis the effect size was the difference between the means of two groups divided by the standard deviation common to the two populations.

A majority of the studies the researchers examined indicated that student performance on examinations was better with visual-based teaching than with traditional teaching methods. Hence, researchers concluded that students learned more from visual-based teaching. The study's results also indicated that students displayed similar attitudes towards classes which used either visual-based teaching or traditional teaching. Researchers found no difference in the two teaching methods in the correlation between student aptitude and achievement, as aptitude played a strong role in student performance in both environments. Finally, students were equally likely to complete courses using either visual-based or traditional teaching.[32]

Studies of Training Media

The studies I have described thus far concern the effectiveness of educational technologies in learning environments. A related field of research deals with evaluating the influence of different training media. Researchers have used meta-analysis techniques to investigate numerous studies to compare the relative achievements of groups of trainees who receive similar training from different media. Effect size is usually defined as the average percent increase in the standard deviations of such outcome measures as examination scores. Evidence from the studies show media as vehicles to deliver instruction but not as vehicles to influence achievement. For example, whenever the same instructor or team of instructors develops both a traditional training program and a program using media, researchers notice no student achievement differences between the programs. However, if different instructors develop the presentations, student achievement differs significantly. Further, the novelty phenomenon causes any advantages from the training media to diminish over time. Researchers cannot conclude whether using traditional training methods or training media produces better

results. In fact, the meta-analysis results indicate that it is the training method and curriculum content, not the medium instructors use to train students, that influences student achievement.[33]

Other meta-analysis reviews, however, conclude that training media do influence student achievement. One such review analyzed studies that compare interactive videodisc instruction with traditional training methods. These studies examined military and industry training and higher education programs. Twenty-four studies compared two different training environments used in military maintenance, operator, and command training programs. The meta-analysis showed that interactive videodisc instruction increased student achievement by 0.39 standard deviations, or an increase from the 50th to the 65th percentile of achievement, when compared to traditional instruction. Researchers analyzed nine studies of industrial training programs, and the results indicate interactive videodisc instruction increased student achievement by 0.51 standard deviations, indicating an improvement in achievement from the 50th to the 69th percentile. Finally, meta-analysis results of 14 studies of instruction in colleges and universities showed that interactive videodisc instruction increased achievement by an average of 0.69 standard deviations, which equate to increasing student achievement from the 50th to the 75th percentile.[34]

This meta-analysis also showed that interactive videodisc instruction was more effective in increasing student achievement than computer-based instruction without videodisc interaction. While interactive videodisc instruction increased student achievement in college courses by 0.69 standard deviations when compared to traditional instruction, computer-based instruction increased achievement by only a 0.26 standard deviations. Conversely, using computer-based instruction in adult education programs resulted in a 0.42 standard deviations increase in student achievement. Still, the meta-analysis results indicated that interactive videodisc and computer-based instructions were more effective than traditional training methods.[35]

Even though researchers may disagree on the effectiveness of media on training programs, they do believe that media can make certain kinds of training more efficient and cost effective. Computers, for example, permit complex interactions which were previously too expensive for regular use; media use permits relatively inexpensive and repeatable training. Media use also decreases the time required for training and tailors instruction to individual trainees.[36]

A Synopsis of the Studies

While box-score reviews and meta-analysis techniques have synthesized many studies dealing with the effectiveness of innovative technologies in education and training environments, one common theme pervades: researchers have found no significant difference between media in facilitating

the achievement of a wide range of learning objectives. In fact, "Most objectives may be attained through [teaching] presented by any of a variety of different media."[37] Instructors can use most types of media effectively to present information necessary to achieve almost any learning objective. They can use printed text and illustrations to teach almost any subject matter to literate learners; the primary issues in using innovative educational technologies include the efficiency of the learning process and the congruency of certain media characteristics with respect to specific objectives and learning tasks. For example, a study discussed earlier in this chapter concluded that pictorial media were suited to presenting concrete information while print media were better for presenting abstractions.[38]

A plausible explanation for the lack of decisive conclusions concerning media effectiveness results from, as Andrew Trotter claims, "The lack of reliable data on educational effectiveness. Hundreds of studies attribute gains in standardized test scores and other improvements to the use of integrated [educational] systems. But many of those studies don't stand up to methodological analysis."[39] Many studies that compare the effectiveness of different media have not clearly defined what they are comparing. While the goal of determining the best technology for teaching is straightforward, it leads to uninterpretable results. In 1963, for example, 250 studies compared teaching using television to face-to-face teaching, yet only 10 were conclusive; the others suffered from a variety of theoretical and methodological inadequacies that stemmed from the researchers' failure to establish concrete study goals.[40]

Coupled with the lack of clear definitions and goals, studies on the effectiveness of media often fail to consider learner variables which interact with the media to bias research results. Differences in such student communication skills and abilities as reading ability, visual literacy, and computer expertise can significantly effect classroom and examination performances. Similarly, cognitive factors like intelligence quotients and learning rates, and personality characteristics like maturity and responsibility, influence student performance in the classroom environment and, consequently, affect the outcomes of the research studies.[41]

Accounting for these variables requires that researchers establish a well-controlled experiment in which they hold constant all possible variables, except for the medium they plan to use in the classroom. The conditions in the learning environments they are comparing are therefore identical, and only the medium of presentation varies. Then the only possible conclusion to the study would be that the media alone would cause differences in student performance. However, Gavriel Salomon and Richard Clark contend that "such differences were rarely found . . . when only the least significant aspects of [teaching were] allowed to vary, nothing of interest could, and did, result."[42] The constraints researchers place on the variables tend to force the different methods of presentation into such similar formats that they can find no significant differences between the achievements of the students exposed to the different presentations.[43]

As for other research efforts, most "media comparison studies clearly suggest that media do not influence learning under any conditions."[44] Media are merely vehicles to aid teaching and they usually do not influence student performance; the capabilities of media negligibly impact student learning from the media. The content of the media influences achievement rather than differences in the media. Even the few studies which demonstrate significant changes in student performance or ability attribute the changes not to the media but to the curricular reform instituted by introducing the media. The effectiveness of innovative educational technologies apparently depends on the teaching style instructors use rather than a particular medium's technological capabilities.[45] Hence, the teaching styles instructors use in different learning environments confound the results of research on media effectiveness.

Uncontrolled novelty effects also confound these results. Students tend to devote increased effort and attention to media that are novel to them; this artificially elevated motivation results from novelty rather than the capabilities of the media. Increased attention may generate achievement gains over traditional teaching methods, but the gains tend to diminish as the students become more familiar with the medium. Recent studies on innovative educational technologies indicate decreased novelty effects when compared to earlier studies; while the media exposure was a new experience for most students in the earlier studies, students today have been receiving substantial exposure to various technologies throughout their education. Hence, students have become less susceptible to the influence of the novelty effect due to their increased exposure to and experience with innovative educational technologies.[46] Still, the novelty effect may distort research results which indicate that teaching using innovative educational technologies is superior to traditional teaching methods.

Alternative Research Directions

Other effects confound the results of research on the effectiveness of innovative educational technologies. Examining the results of studies of research in this field and considering the myriad of confounding factors leads to Richard Clark's caution:

> Based on this consistent evidence, it seems reasonable to advise strongly against future media comparison research. Five decades of research suggest that there are no learning benefits to be gained from employing different media in [teaching], regardless of their obviously attractive features or advertised superiority. All existing surveys of this research indicated that confounding has contributed to the studies attributing learning benefits to one medium over another and that the great majority of these comparison studies clearly indicate no significant differences. . . . Where learning benefits are at issue, therefore, it is the method, aptitude, and task variables of [teaching] that should be investigated. Studies comparing the relative achievement advantages of one medium over another will inevitably confound medium with method of [teaching].[47]

Studies Examining Media Influences on Learning

Predicated on this supposition, researchers who study innovative educational technologies have shifted their focus in recent years. One aspect of educational technology researchers have examined concerns whether such media attributes as varying difficulty levels, entertainment value, and enjoyment might influence student achievement in a learning environment. Studies indicate that "student beliefs about the different demands placed on them by different media influenced their approach to learning tasks."[48] Student perceptions of the medium and their own abilities directly relate to the effort they invest in the learning process rather than in the medium itself. For example, students typically perceive television as an easy medium from which to learn and attribute great difficulty to learning from computers.[49]

Similar studies have examined the relationship between the media and student enjoyment. Students who were exposed to computer-assisted instruction liked the computer less but learned more from it than from other media. In addition, while students liked television less than voice recordings as a learning medium, they learned significantly more from the television. Researchers have suggested that students preferred learning from media that inadvertently induced less learning for those students. While higher ability students preferred more structured and directive media because they thought they would have to invest less effort to achieve success, they could not optimize their skills and consequently achieved less than when they used less directive media. Conversely, students of lower ability liked less structured and more discovery-oriented media to avoid investing the effort required by the structured media to achieve the same substandard results. However, these students needed the structured format they disliked to improve their performance.[50]

Research also has focused on examining how various media attributes influence the learning process when they interact with the student's cognitive processes. The ways students internally process different presentation modes and how they develop these processing capabilities influences learning heavily. As an example, studies have shown "that children attend to televised material that is comprehensible to them, implying that comprehensibility determines attention rather than the other way around."[51] Educators should therefore develop educational media to convey comprehensible information rather than merely to attract attention. This focus in media research identifies media attributes which distinguish between various media while affecting cognitions relevant to the learning process. Instead of asking which medium teaches better, research studies identify which media attributes might combine with student cognitive traits under different conditions to produce the desired kinds of learning.[52]

Thus, the research encompasses more than merely finding the "best" style of teaching or the "best" innovative educational technology. Rather, studies can identify situations in which a particular style of teaching or medium enhances the learning process, those situations in which there is no difference in the learning process, and those in which such an approach or medium is inappropriate.[53]

Researchers in this field have advocated several orientations to provide useful direction for experimental research. One promising orientation emphasizes the external validity of studies, which is the degree to which the study and the corresponding results relate to real-life applications. Recall that internal validity is the extent to which researchers eliminate external variables which influence the learning process in a classroom, lending a sterile atmosphere and permitting researchers to conduct a controlled study. Previous studies have established high internal validity by eliminating sources of extraneous variance to assess student performance in different learning environments. Minimizing the degree to which environmental variables and student and instructor characteristics influence the research increases the experiment's internal validity. While this heightened internal validity enhances conceptualization and understanding of the study results, conclusions obtained from contrived or artificial applications may have little generalized implications and, hence, only remote relevance to realistic educational practices. Unfortunately, studies with real-world applications deal with complex variables and by nature are highly specific. The complexity results in poor internal validity, while the study's narrow focus precludes generalization to associated educational technologies. Consequently, in this area of research, studies emphasize a high degree of internal validity while concentrating on maximizing the degree to which educators can relate the results to real-life applications.[54]

Studies of Media Replications

A second research orientation concentrates on media replications rather than on media comparisons. The latter focuses the research on the media while it attempts to control the teaching styles instructors use in the learning environment. As previously mentioned, controlling the confounding variables associated with this research is impractical. More important is whether comparing media makes any sense given that they are just conduits.[55] As Richard Clark states, "It seems reasonable to recommend, therefore, that researchers refrain from producing additional studies exploring the relationship between media and learning unless a novel theory is suggested."[56] In contrast, media replications test the reliability and practicality of teaching styles using different delivery modes. Comparing the performance of students exposed to various teaching styles across the spectrum of innovative educational technologies can help to understand these styles further and to identify ways to apply them in the learning environment more effectively.[57] Future research could focus on characteristics of teaching styles and such variables as task and learner aptitude, which can be effective means for understanding increases in student performance.[58] While studies of this orientation examine the teaching styles themselves, insight into the presentation modes may sometimes occur.

An example of such research is a series of studies conducted in the mid-1980s that "examined the effectiveness of personalizing math word

problems by embedding information about the individual student in the problem context."[59] Instructors either generated and presented the problems to the students using computer-assisted instruction or had the computer generate and present them in print form. Both delivery systems showed this style of drill-and-practice instruction to be more effective than conventional ways of presenting math word problems. The media replication study also discovered differences in the two delivery modes. Specifically, computer-assisted instruction simplified administration and logistics, in addition to providing greater experimental control by preventing students from skipping around to either review previous material or see how much material remained. Students also worked significantly longer and seemed more enthusiastic about using computer-assisted instruction. Consequently, these studies not only concluded that either delivery application supported the strategy of personalizing instruction but also suggested using computer-assisted instruction rather than print as the delivery mechanism. Note that while media replication appears similar to media comparison, the orientation of each is diametrically opposed; replication uses media to assess teaching styles rather than using the styles to assess media.[60]

A study conducted in 1989 also focused on computer-assisted instruction, but researchers designed this study to compare the performance and attitudes of students working individually on a computer-based sex education lesson with those of students working together in dyads on the same lesson. The researchers divided 60 randomly selected eighth-grade students into two groups and either assigned partners or tasked them to work independently on the lesson. They then compared the two groups by examining the results of tests instructors administered after the students completed their lesson. The students also completed a survey to assess their attitudes toward both the instruction and the lesson content.

The study results indicated that students working cooperatively performed significantly better than those working individually. Students with high ability indicated no preference on the instructional method and generally had positive attitudes towards both methods. In contrast, low-ability males preferred individualized instruction over the cooperative method, while low-ability females preferred cooperative instruction. Finally, the researchers noted several variables which might have biased the study results: the short-term nature of the study might have influenced group performance; the durability of the results was questionable, since they were based on an immediate posttest; and the results may have been affected by such factors as the novelty of the lesson, the methods of instruction the instructors used, and the sensitivity of the lesson content.[61]

Studies of Learner Control Variables

A controversial orientation towards media research uses learner control to adapt validated teaching styles to individuals. As educational technology

continually expands, some researchers regard the medium as the agent that acts directly on learning and cognition, not the teacher or lesson.

The medium controls learning and makes decisions for the teacher and students. Unfortunately, this discourages educators from adapting teaching to account for individual differences. The learner has little control over the learning process. Allowing students greater autonomy in making decisions about their own learning increases their learning and motivation significantly. Research therefore identifies learner control variables which are effective for different learners and tasks.[62]

Studies of the Utility of Educational Technology

A final research orientation has distinct ethical implications, since it examines the utility of innovative educational technologies. Schools tend to adopt educational technologies in their curriculum in response to external pressures from special interest groups rather than from an identified and expressed educational need.

Further, it appears that most new media do not focus primarily on educational applications; their effective use in the learning environment is a beneficial spin-off for these media. Hence, according to Richard Clark and Gavriel Salomon, "In the future, researchers might ask not only how and why a medium operates in [teaching] and learning, but also why it should be used at all."[63]

Summary

The published studies I have summarized in this chapter indicate that, in general, researchers have found little difference between traditional teaching methods and teaching using educational technologies. In fact, as Richard Clark emphasizes, "There are no learning benefits to be gained from employing different media in [teaching], regardless of their obviously attractive features or advertised superiority."[64] Results invariably indicate insignificant differences between various presentation modes, whether they use innovative educational technologies or traditional teaching methods.

Researchers do, however, have available a plethora of related research topics which can benefit the field of education. For example, they can compare the effectiveness of one specific traditional teaching method to another to determine which is more appropriate for certain courses or educational disciplines. Similarly, researchers can investigate specific innovative educational technologies and examine them to determine for which types of courses or disciplines they are better suited. Educators require such studies of teaching styles if they wish to develop the optimum learning environment for their courses and students.

The Air Force Academy is uniquely suited to conduct controlled experiments on teaching styles and their effectiveness in increasing or enhancing student

learning. Studies researchers conduct with the networked classroom system can possess both high internal and external validities. Researchers can either control or account for many of the variables attributed to the internal validity of a research study at the Air Force Academy. In addition, instructors can develop courses which emphasize real-life applications, thereby increasing the external validity of the studies. The Civil Engineering Department's Air Base Planning Course is one such course; it engages students in building and maintaining air bases when they are faced with various realistic scenarios. Studies which possess high internal and external validity also can focus on "learning environment replications" by examining the relationship between various teaching styles and delivery modes. The Air Force Academy can generate such world-class research studies.

A potential research effort closely related to comparing various teaching styles examines the appropriateness of these teaching styles for different educational disciplines. Existing studies usually fail to discriminate between the types of courses researchers use for the experiment. Instructors may better teach courses in different disciplines using specific teaching styles or innovative educational technologies. For example, a discussion teaching style may work well in such humanities courses as history and English literature, but it is usually not appropriate for science courses like calculus and physics. One aspect of this research effort seeks to determine if educators can generalize the conclusions of previous studies for all disciplines, or if they are discipline-specific. Educators also can conduct studies to determine the usefulness of different teaching styles or innovative educational technologies for specific courses within different disciplines.

Before educators can accomplish any of these potential research efforts, they must first examine methodologies which they can use to assess the effectiveness of different teaching styles. The next chapter summarizes some of those methodologies.

Notes

1. Gavriel Salomon and Richard Clark, "Reexamining the Methodology of Research on Media and Technology in Education," *Review of Educational Research* 47, no. 1 (Winter 1977): 99–100.
2. Ibid., 105.
3. Ibid.
4. Ibid., 106–7.
5. Ibid., 107.
6. W. Howard Levie and Kenneth E. Dickie, "The Analysis and Application of Media," in *Second Handbook of Research on Teaching*, ed. Robert M. W. Travers (Chicago: Rand McNally College Publishing Company, 1973), 860.
7. Ibid.
8. Ibid., 863–66.
9. Ibid., 867–73.
10. Peter A. Cohen, Barbara J. Ebeling, and James A. Kulik, "A Meta-Analysis of Outcome Studies of Visual-Based Instruction," *Educational Communication and Technology Journal* 29, no. 1 (Spring 1981): 27.

11. Simon Hooper and Michael J. Hannafin, "Psychological Perspectives on Emerging Instructional Technologies: A Critical Analysis," *Educational Psychologist* 26, no. 1 (1991): 80.

12. Benny M. White, Kenneth M. Matthews, and C. Thomas Holmes, "The Cost-Effectiveness of an Interactive Video System for Science Instruction," *Educational Technology* 29, no. 12 (December 1989): 19–23.

13. David L. Silvernail and Judith L. Johnson, "The Impact of Interactive Televised Instruction on Student Evaluations of Their Instructors," *Educational Technology* 32, no. 6 (June 1992): 48.

14. Ibid., 47–50.

15. Ibid., 48–50.

16. Robert B. Kozma and Jerome Johnston, "The Technological Revolution Comes to the Classroom," *Change* 23, no. 1 (January/February 1991): 14–15.

17. David W. Dalton and Michael J. Hannafin, "The Effects of Computer-Assisted and Traditional Mastery Methods on Computation Accuracy and Attitudes," *The Journal of Educational Research* 82, no. 1 (September/October 1988): 27–33.

18. Larry F. Fine, R. C. Bialozor, and T. F. McLaughlin, "An Analysis of Computer-Assisted Instruction on Scholastic Aptitude Test Performance of Rural High School Students," *Education* 111, no. 3 (Spring 1991): 400–403.

19. Omer Geban, Petek Askar, and Ilker Ozkan, "Effects of Computer Simulations and Problem-Solving Approaches on High School Students," *The Journal of Educational Research* 86, no. 1 (September/October 1992): 5–10.

20. Ibid.

21. Ibid.

22. Ibid.

23. James A. Kulik, Robert L. Bangert-Drowns, and George W. Williams, "Effects of Computer-Based Teaching on Secondary School Students," *Journal of Educational Psychology* 75, no. 1 (February 1983): 19.

24. Ibid., 20.

25. Cohen, Ebeling, and Kulik, 27–30.

26. Kulik, Bangert-Drowns, and Williams, 19.

27. Ibid., 20–21.

28. Ibid., 21–22.

29. Ibid., 22.

30. Ibid., 22–26.

31. Richard Clark and Gavriel Salomon, "Media in Teaching," in *Handbook of Research on Teaching*, 3d ed., ed. Merlin C. Wittrock (New York: Macmillan Publishing Company, 1986), 466.

32. Cohen, Ebeling, and Kulik, 26–36.

33. Richard Clark, "The Role of Media in the Evaluation of Training," in *Evaluating Business and Industry Training*, ed. Leslie S. May, Carol A. Moore, and Stephen J. Zammet (Boston: Kluwer Academic Publishers, 1987), 153–58.

34. J. Dexter Fletcher, *Effectiveness and Cost of Interactive Videodisc Instruction in Defense Training and Education*, IDA Paper P-2372 (Alexandria, Va.: Institute for Defense Analyses, July 1990), s-2-IV-6.

35. Ibid.

36. Clark, "The Role of Media in the Evaluation of Training," 153–58.

37. Levie and Dickie, 859.

38. Ibid.

39. Andrew Trotter, "Computer Learning with Integrated Systems," *The Education Digest* 56, no. 4 (December 1990): 29.

40. Levie and Dickie, 860.

41. Ibid., 859.

42. Salomon and Clark, "Reexamining the Methodology of Research on Media and Technology in Education," 101–2.

43. Clark and Salomon, "Media in Teaching," 465.

44. Richard Clark, "Reconsidering Research on Learning From Media," *Review of Educational Research* 53, no. 4 (Winter 1983): 445.

45. Ibid., 445–46; Hooper and Hannafin, 70; Clark and Salomon, "Media in Teaching," 466.

46. Dalton and Hannafin, 32; Clark, "Reconsidering Research on Learning From Media," 449–50.

47. Clark, "Reconsidering Research on Learning From Media," 450.

48. Ibid., 455.

49. Ibid.

50. Ibid., 455–56.

51. Clark and Salomon, "Media in Teaching," 467.

52. Ibid., 467–73.

53. Amy Rose, "Technology—Servant or Master?" *Adult Learning* 4, no. 3 (January/February 1993): 6.

54. Salomon and Clark, "Reexamining the Methodology of Research on Media and Technology in Education," 107–8; Steven M. Ross and Gary R. Morrison, "In Search of a Happy Medium in Instructional Technology Research: Issues Concerning External Validity, Media Replications, and Learner Control," *Educational Technology Research and Development* 37, no. 1 (1989): 21–30.

55. Ross and Morrison, 26.

56. Clark, "Reconsidering Research on Learning From Media," 457.

57. Ross and Morrison, 26–30.

58. Clark, "Reconsidering Research on Learning From Media," 457.

59. Ross and Morrison, 27.

60. Ibid.

61. David W. Dalton, Michael J. Hannafin, and Simon Hooper, "Effects of Individual and Cooperative Computer-Assisted Instruction on Student Performance and Attitudes," *Educational Technology Research and Development* 37, no. 2 (1989): 15–24.

62. Ibid., 28–30.

63. Clark and Salomon, "Media in Teaching," 475.

64. Clark, "Reconsidering Research on Learning From Media," 450.

Chapter 4

Assessment Methodologies

At first glance, assessing the effectiveness of different teaching styles appears to be straightforward; a style of teaching is effective if it improves a student's performance. Unfortunately, this simplistic view belies the complexity of assessing the effectiveness of teaching styles and fails to account for numerous important considerations. Comprehensively assessing student performance is a demanding and time-consuming process, but it "lies at the heart of successful teaching."[1] Therefore, before reviewing existing assessment methodologies, I must first discuss factors which are relevant to the assessment process.

Considerations in Developing and Using Assessment Methodologies

A logical starting point in reviewing these assessment methodologies is to define them first. According to K. Patricia Cross and Thomas Angelo, assessment methodologies are instruments and methods which inform instructors and evaluators about what students learn in the classroom and how well they learn.[2] Assessments involve using a variety of instruments to gather information, and they are most meaningful when integrated into the teaching style and learning process.[3] To develop a sound educational assessment, educators must

1. choose a behavior or cognition of significant interest,
2. clearly specify a task to elicit the behavior or cognition,
3. observe the relevant portion of the behavior or cognition, and
4. interpret the demonstrated behavior or cognition.[4]

The result of any assessment is a decision; information is the primary output of an assessment and the primary input to a decision making process.[5] These decisions range from assigning grades to students to deciding the future status of instructors or learning environments. The decision makers' values, priorities, and attitudes influence an assessment and their actions resulting from that assessment. Educators must therefore consider the decision makers' agendas when designing an assessment.[6] The following nine basic principles help guide the educators in developing assessment methodologies.

1. The assessment should be authentic and meaningful so students can effectively use their knowledge and existing resources to achieve a desired effect.

2. Design performances rather than drills to provoke thought.

3. Specific tasks in the assessment should be valid samples addressing course objectives, with sufficient depth and breadth to permit generalizations about overall student performance.

4. The assessment tasks should be rich, realistic, and enticing while considering available time and resources.

5. The tasks should be validated to justify their use.

6. Scoring criteria should be authentic by recognizing essential successes or errors in understanding the material, not by scoring what is easy to count or observe.

7. The performance standards associated with the scoring should be genuine benchmarks, not arbitrary scores.

8. The scoring should be realistic and reliable, relying on descriptive language instead of evaluative or comparative language.

9. Educators should report and use the assessment results to insure that achievement and progress are related to essential performance and established standards.[7]

Validity, reliability, and equity are three principle standards educators must adhere to as they develop an assessment. Validity relates to the appropriateness, meaningfulness, and usefulness of the assessment; it measures what educators truly value. Reliability not only refers to how test results match other tests, but how consistent the scoring is and how well different educators agree with the results. Equity can be viewed as the elimination of bias.[8] The effectiveness of an assessment depends on how well educators enforce these three standards.

Determining Student Performance

Educators who design assessments also must consider that human learning involves multiple processes, each of which obeys somewhat different laws and has different implications for teaching. Students differ in the knowledge, skills, abilities, and understanding each brings to the learning environment. Their individual characteristics are among the most important determiners of learning outcomes. Preexisting ability levels and differences among students in aptitudes, motivation, and interest strongly influence the rate that students acquire knowledge and understanding, as well as their performance in learning environments. In fact, such student characteristics as aptitude, reading, and motivation levels are some of the most powerful determinants of student performance. These variables impact performance more than such course-content variables as course materials and depth of information covered; the most effective type of course content manipulates a variety of student characteristics.[9]

In fact, content validity, or the degree the course content reflects each student's critical knowledge, skills, and abilities (KSA) that are required for satisfactory performance, determines if the content is valid or how it should be changed. Establishing critical KSAs requires extensive knowledge of the course objectives and goals.[10] In general, three broad goals exist for courses: know important facts and constructs; gain skills and abilities appropriate for the discipline; and acquire the dispositions to apply the learned knowledge, skills, and abilities to new situations.[11] Student knowledge is the foundation upon which skills and abilities are built; it is an organized body of information of a factual or procedural nature which makes adequate performance in the classroom possible. However, possessing knowledge does not insure its effective use. Skills and abilities put knowledge into practice. Skills refer to the capabilities of students to perform in a learning environment, while abilities are cognitive capabilities necessary to accomplish specific functions within the learning environment.[12]

Educators must specify and relate KSAs to the objectives and goals of the course. Otherwise, the KSAs they emphasize in the course may not relate to their goals and objectives, or they may not specify KSAs which could relate to the course goals and objectives. KSAs describe the characteristics of good and bad performances and define what students need to know to be successful in class.[13]

Assessing the students' KSAs is a systematic process, and educators must develop a working group to develop and accomplish these assessments; decision makers, instructors, evaluators, and information users usually comprise this group. First, the decision makers identify the need for the assessment. They then help the instructors and evaluators develop a strategy to implement the assessment. The instructors teach the course the working group is assessing. Evaluators collect and analyze data on the students and the learning environment to provide the most accurate and useful information for decision making. Finally, the evaluators communicate the assessment results to the decision makers and the information users in a timely and efficient manner.[14] Throughout this assessment process, the decision makers must specify what they wish to evaluate. Determining student performance requires answers to such basic questions as:

1. With respect to student performance, what questions should educators answer? What should students know and what should they be able to do?
2. What can evaluators measure to answer those questions?
3. What dimensions of learning or performance are evaluators measuring?
4. What criteria will count as satisfactory performances?
5. What sources of data can evaluators use to help measure performance?
6. What alternative ways of gathering data exist?
7. What assessment criteria should educators apply to answer each question?
8. How can evaluators assure expert and unbiased judgments of the quality of a student's work?
9. How can educators provide feedback?[15]

Categories of Assessments

These types of questions allow evaluators to assess student performance in various ways; they can employ eight categories of assessments:

1. system analysis: quantitatively measures program inputs and outcomes to examine effectiveness and efficiency
2. behavioral objectives approach: focuses entirely on clear, specific, and measurable goals
3. goal-free assessment: examines the extent to which the course meets actual student needs
4. art criticism approach: makes the evaluator's own experience-derived standards of excellence a criterion against which performance is judged
5. accreditation model: uses teams of external accreditors to determine the extent to which a program meets professional standards for a given type of course
6. adversary approach: permits two teams to battle over the summative question of whether educators should continue a course or teaching style
7. transactional approach: concentrates on the progress of the course to determine student performance
8. user-focused assessment: focuses on and is driven by the information requirements of specific people who will use the assessment findings.[16]

While these categories are fairly comprehensive, assessments the working group derives from the categories must be flexible to adapt to changing learning environments and to different people and situations. Creative evaluators use different assessment methodologies and match them to unique situations.[17] Students tend to learn according to how evaluators assess them. The student performances evaluators witness are those they test for; if they don't assess it, they probably won't observe the performance. If an evaluator tests students solely for facts, they will memorize facts. If he or she has them to analyze situations, the students learn to think critically.[18] Therefore, in selecting the appropriate assessment methodologies, the working group must ask such probing questions as:

1. Will the assessment methodology provide information about what students are learning in individual classrooms?
2. Does the methodology focus on aspects or variables of learner and instructor behavior or teaching style characteristics that can be changed to enhance learning?
3. Will the methodology provide teachers and students information they can use to make midcourse changes and corrections?
4. Is the assessment methodology relatively simple to prepare and use?
5. Are the results acquired relatively quickly, and are they easy to analyze?[19]

In addition to asking these questions, the working group must determine the goals for student achievement. According to Richard Stiggins, "Four broad categories of achievement targets are: (1) substantive subject-matter

knowledge to be demonstrated, (2) thinking skills to be demonstrated, (3) specific desired behaviors to be exhibited, and (4) products with specific attributes to be created."[20] To decide if students actually achieve these goals, evaluators can use one of three general methodologies: "paper-and-pencil" assessments, performance assessments based on observation and judgment, and direct personal communication with the student.[21] Each assessment category previously described relies on one or more of these methodologies. When used with an appropriate achievement goal, the methodology provides legitimate data for decision makers.

Models of Assessments

Evaluators do not use these three methodologies independently; rather, they employ various combinations of them. Still, evaluators use three basic models to categorize most assessments. The procedural model assesses the process of transmitting knowledge to the student instead of assessing the student's level of knowledge itself. This model forces the achievement goals to be subjective by not specifying the kind or degree of knowledge, skill, or ability required of the student. Evaluators do not assess teaching by a common standard, and they emphasize fairness, due process, and objectivity in making judgments of student performance. The evaluators study such materials as syllabi, course notes, and examinations as part of the evaluation process. A second model, the quantitative-mathematical model, emphasizes criteria, standards, and information while de-emphasizing the importance of the learning process. Evaluators compile many forms of data, such as student feedback forms, classroom visitations, and examination materials. They state standards in quantitative terms, and they convey judgments on student performance through mathematical formulas. The final model, the learning outcomes model, judges effectiveness by whether a student has learned. This model uses norm- or criterion-reference measurements of learning and emphasizes the purpose, outcome, and impact of the learning process on the student. It uses pre- and posttests extensively and de-emphasizes student feedback forms. The learning outcomes model pays little attention to the process of learning, and instead focuses on performance criteria and standards.[22]

Measuring Student Performance

These models differ in their measurement techniques to determine if students have achieved the established goals and objectives of the course; however, they provide tools for educators to develop methodologies to assess the effectiveness of teaching styles. The ultimate criteria for acceptable student performance for any of the models is whether the student has

accomplished a measurable gain toward specific course goals and objectives. Unfortunately, the problem that arises in using student gains to measure the effectiveness of a particular style of teaching is determining the types of gains to measure. Another problem is contamination; that is, the gains may not be solely due to the style of teaching. A student's general mental ability, past educational experience, peer and instructor influences, and extracurricular activities influence the learning process.[23]

Given these problems, measuring performance based on student gain poses a challenging prospect. Quite often evaluators quantify student performance in terms of gain scores extracted from testing. They administer a pretest to students at the onset of a course and a posttest at the conclusion of the course. Gain scores are then explicit differences between the pre- and posttest scores and indicate student achievement during the course. The simplest type of gain score is the difference score, which is the difference between pre- and posttest scores. Typically, this score is unreliable due to the usually high correlation between the pre- and posttest scores; for example, students with high posttest scores also have high pretest scores. Further, since the difference score is directly related to the pretest score, evaluators find it difficult to make adjustments for the individual differences of students; therefore, they cannot compare individuals or groups with unequal pretest scores. When evaluators do compare two groups with different pretest scores, they note that the group with the lower pretest scores tends to have larger gain scores than the other group. Difference scores cannot, in general, account for the nonequivalence between groups.[24]

Evaluators employ alternative methods to difference scores to determine student achievement in the classroom. For example, they can use residualized change scores, which are gains uncorrelated to the pretest scores. Evaluators predict a student's posttest score by using the linear regression of the posttest on the pretest and then comparing the student's actual posttest score to the predicted one. The evaluators quantify a student's performance based on whether his or her posttest score is more or less than the predicted score. Unfortunately, residualized change scores do not correspond to measures of individual growth curves and are not true measures of change.[25]

Estimates of true change, on the other hand, reflect corrected measures of change. Evaluators use regression analysis and classical test theory assumptions of true scores and error scores that underlie the observed pre- and posttest scores to measure true change. This method requires reliability estimates of both the pretest and the posttest scores, which are not always available.[26]

Another alternative to the difference scores is the standardized change scores, which are simply standardized difference scores. Specifically, evaluators set the variances of the pre- and posttest scores to be equal, and they can then compare groups or individuals with unequal pretest scores. However, this scoring system assumes the students in the comparison groups are similar in age, race, socioeconomic status, and other characteristics which may influence student performance.[27]

Repeated measures analysis of variance (ANOVA), a statistical method for analyzing assessment data, is also similar to difference scores. With both methods evaluators assume a perfect relation between the pre- and posttest scores; ANOVA contrasts with difference scores in that in addition to analyzing pre- and posttest scores, it also examines the relationship between two variables in the data. Evaluators represent the first variable as a subjective variance between the experimental and control groups; the second variable is a time factor. Evaluators generate these variances by measuring such subjective variables as student satisfaction or attitude. The degree of precision of ANOVA depends on the degree of correlation between the pre- and posttest scores.[28]

Evaluators can account for this correlation by using the analysis of covariance (ANCOVA). With this method, they assign a weight to the pretest scores to account for the correlation between the pre- and posttest scores. The evaluator can then control the effect of the pretest while examining the difference between the experimental and control groups on the posttest. However, if there is little or no correlation between the pre- and posttest scores, ANCOVA may be less than optimal.[29]

While evaluators can use ANOVA or ANCOVA to examine the effects of different independent subjective variables on the outcome of the assessment, they complicate the assessment of the change in a student's performance by introducing these variables in the study. Evaluators can measure the variables by using surveys and questionnaires students complete both at the beginning and the end of a course; a student's scores on the questionnaires reflect different impressions at the beginning of the course when compared to the end of the course. One way to account for these phenomena is to have the student complete a retrospective questionnaire at the end of the course to obtain student impressions of the course retrospectively; that is, before the course began.[30] Another problem with assessments using ANOVA and ANCOVA, as well as the other gain scores methods, is that evaluators can use them to examine only the outcomes of the course. Assessments based on having evaluators test students prior to the beginning of a course and after completing the course lack critical information about what happens during the course.[31]

David Dalton, Michael Hannafin, and Simon Hooper employed a different method to measure student performance when they compared students working individually and in dyads. Before these researchers began the study, they grouped the students according to their scores on the Comprehensive Test of Basic Skills (CTBS), a norm-referenced measure of general achievement that provides a global estimate of a student's overall ability. The CTBS permitted the researchers to group the students as either "high" or "low" in learning ability, and they randomly assigned the students, within gender and ability, to two groups; one group worked individually on an assignment while the other group worked in pairs. Instructors then administered a posttest to the students in both groups after they completed the assignment, and the evaluators compared the means of both groups' test

scores.[32] Identifying high- and low-ability students facilitated comparing high-ability students who worked individually with high-ability students who worked in dyads. Similarly, the evaluators could compare the performance of low-ability students in the two groups. Dalton, Hannafin, and Hooper assumed the process of randomly assigning blocks of high- and low-ability students to the two groups to equalize the groups. In other words, assuming random differences between individuals, the two blocks of high-ability students were equal in ability; similarly, the two blocks of low-ability students were equal in ability.[33] However, evaluators did not administer pretests to determine differences in knowledge and experience.

Note that David Dalton and Michael Hannafin also used this measurement technique to compare traditional teacher-centered teaching with computer-assisted instruction in mastery learning. These researchers once again used the CTBS to identify high- and low-ability students and then randomly assigned them to the comparison groups. After the students completed their assignments, they completed posttests, which evaluators used to compare the groups' performances against each other.[34] As before, the evaluators did not administer pretests, nor did they compare the posttest results to any baseline tests; Dalton and Hannafin assumed that the blocks of high- and low-ability students were equal in knowledge and experience prior to the assignment.

Educators have developed additional methodologies which use tests to measure student performance, and the utility of these methodologies depends on the users of the information. One consideration in using test scores is the form in which evaluators present them. Such raw scores as the number of items a student correctly answers usually have little meaning by themselves, since evaluators cannot generalize them to define achievement domains clearly. They must transform the scores to forms that can provide information about a student's standing relative to a norm group. Grade equivalents, percentile rankings, T-scores, and normal curve equivalents are examples of techniques which transform the raw scores to useful forms.[35] Each of these techniques has advantages and disadvantages, and evaluators must choose one that best satisfies their analyses.

Types of Tests

Before evaluators can decide which technique to use to assess student performance based on test results, they must first consider the content of the tests themselves. Tests determine what instructors teach and what students study. When educators specify learning objectives, they invariably design tests to be isolated exercises which elicit the desired objectives. Instructors teach the precise content of the tests instead of underlying concepts, and they teach skills in the same format as the tests rather than how students would use the skills in the real world. Fair and valid tests should not provide blatant cues and demand simple recall of material; rather, they should center around

the students' thought processes by requiring more complex and challenging mental processes. Acknowledging that there is usually more than one approach to a problem or situation, educators should construct tests that place more emphasis on uncoached explanations and allow students to develop their own cognitive maps relating concepts to facts.[36] Many existing assessments can accomplish this task; for example, some educators have designed norm-referenced assessments to do so.

Norm-Referenced Assessments

Quite often, evaluators rely on such norm-referenced assessments as standardized tests to measure student performance against established criteria. Blaine Worthen states that "for most of the 20th century, standardized, multiple-choice tests have served as the primary method for assessing [student performance]."[37] However, educators have found faults with these tests; for example, instructors may teach the tests, especially if they feel their jobs depend on their students' standardized scores.[38] In addition, these tests limit both the breadth and depth of content coverage. They also invariably emphasize basic skills and do not cover the full range of important behavioral and cognitive objectives. E. Peter Volpe believes that "most knowledgeable educators contend that a standardized examination measures only how well a student has studied for the examination."[39] Students tend to memorize data for factual recall rather than to analyze and synthesize information.

To simplify the process of grading standardized tests, evaluators usually employ multiple-choice formats, which prevent students from explaining their answers.[40] Still, multiple-choice tests extensively sample "course content in a relatively quick and easy manner [and can] measure students' knowledge of specific facts, principles, methods, and procedures, plus the application and integration of data."[41] Evaluators use such tests to

1. determine student responses to who, what, when, or where type of objectives;
2. assess student knowledge of isolated facts, definitions, or vocabulary usage; or
3. determine if students can correctly solve a specific problem.[42]

While multiple-choice tests can elicit important conceptual distinctions, they don't measure the students' abilities to organize relevant information and present coherent arguments. Further, students may legitimately need to have a question rephrased occasionally or have the opportunity to defend "incorrect" answers; to fully assess a student's performance, evaluators must establish a dialogue with the student.[43] On the other hand, they can design multiple choice tests to assess and effectively judge a student's higher order thinking skills and abilities and his or her understanding of the material by, for example, leading the student through the application of a problem or situation.[44]

Unfortunately, a multiple-choice test is an example of a norm-referenced assessment that provides marks to compare a student's performance to the performance of other students but offers little information about the student's achievement. Evaluators who use such norm-referenced assessments assume teaching has the same effect on all students in the course, and the performances of a comparison group determines the quality of a student's performance.[45] Based on these assessments of relative importance, instructors assign students comparative grades which do not indicate how the students compare to a set performance standard.[46] In addition, one essential feature missing from most norm-referenced assessments is constancy; each test created defines a unique scale, and the students' scores are bound to that scale alone. Item banking, however, is a norm-referenced assessment technique which establishes a performance standard and achieves constancy of scale for all tests within a particular course.[47]

Item banking allows evaluators to catalogue a large collection of test questions in terms of content and difficulty. The questions assess one specific area of performance, and evaluators sort them in degrees of difficulty that range from easiest to hardest. They can tag students answering the questions onto a bank scale in terms of performance on the questions; the students and the questions appear on the same scale and are measured in the same units. Thus, evaluators generate tests that are calibrated to a standard bank scale before they are distributed to students. Item banking also provides to evaluators and instructors a means of monitoring a student's progress. By screening the test results of specific questions, instructors can determine if the students have failed to understand what they taught. Further, if an individual misses a question that most students of the same achievement level easily answered, instructors can ask the student additional related questions to assess the individual's strengths and weaknesses in that subject area. Finally, item banking lends itself to computerized tutoring. If a student engaged in a tutor program on a computer misses a question on a particular topic at a certain level of difficulty, the student will view the correct answer on the computer screen, and then the computer will display a series of item banked questions on the same topic at the same difficulty level. Once the student answers the question correctly, the computer will display questions on the same topic at increasing levels of difficulty to insure the student understands the topic.[48]

Another approach to assessing a student's performance with respect to a preestablished standard is absolute grading. Evaluators formulate these standards from the specific course behavioral and cognitive objectives and they equate them to the level of performance required to demonstrate mastery of the topic. Evaluators can identify test items on examinations with specific objectives and score the items in terms of the degree to which students attained the objectives. If a student meets or exceeds the specified level of performance for an objective, he or she demonstrates mastery of that objective and earns a maximum score. However, if the student does not demonstrate mastery by meeting the level of performance, he or she receives a score that is

less than the maximum; evaluators establish a scale of scores for varying levels of performance. Evaluators determine the quality of each student's performance in terms of whether he or she has demonstrated a particular mastery. Hence, the standards are stable for all students regardless of the composition of the class.[49]

Such norm-referenced assessments as item banking and absolute grading provide evaluators with a means to establish performance standards and to monitor student progress. However, too often evaluators administer norm-referenced assessments only a few times during the course, leaving them without the ability to gauge a student's progress over an extended period. While norm-referenced assessments concentrate on a student's short-term recall, the true indicators of a student's ability are whether that student can, over an extended period, perform consistently well tasks whose criteria for success are known and quantified. Standardized tests alone cannot measure student growth. Tests should challenge students to extend their knowledge, to criticize, and to explain and explore the limits and assumptions on which theories and contrived situations rest; such tests should relate to the real-life environment which pertains to the course to determine whether the students will succeed in the academic discipline or the corresponding workplace.[50]

Criterion-Referenced Assessments

Although assessments based on real-life scenarios seem contrived, evaluators can realistically accomplish them. According to Sally Brown, "In 'real life' we constantly assess writing, dancing, painting, scientific activities, historical analyses, and competence in communication, without assigning numbers or grades; nor do we put them in rank order."[51] Criterion-referenced assessments qualitatively describe what a student has achieved without referring to other students' performances, permitting evaluators and instructors to form rational decisions on how effectively the student achieves the lesson objectives. These assessments compare information about student performance to specified and predetermined levels and include knowledge, skills, abilities, and attitudes the student has acquired. Knowledge may be substantive, emphasizing learning outcomes, or experiential, emphasizing learning through reconstructing student experiences. Skills and abilities may be practical or intellectual, and attitudes may be emotional or cognitive.[52]

A common criterion-referenced assessment is the essay test; it allows students to "apply, analyze, synthesize, and evaluate ideas and information."[53] Students construct their responses to questions that allow them the latitude to express their own thought processes without the constraints of the selected answer formats of such norm-referenced assessments as multiple-choice tests. Because essays normally don't limit student insight and creativity, evaluators can use them to measure a student's critical thinking rather than just his or her possession of knowledge.[54]

Another example of a criterion-referenced assessment is the Scottish Certificate of Education, Standard Grade. Educators use this methodology to determine student performance and progress toward attainment target sets for such subjects as English, mathematics, and science.[55] Instructors administer standard assessment tasks (SAT) which are "externally provided tasks and procedures designed to produce performance data on a national scale."[56] This assessment methodology has several elements, or domains, for each academic subject tested; within each domain evaluators can identify specific levels of achievement. For example, evaluators conceptualize English as writing, speaking, reading, and listening. The levels of achievement for, say, writing include: conveying information, constructing and expressing ideas, arguing, evaluating, describing personal experiences, expressing feelings, and employing literary forms. Evaluators assess students against summary grade-related criteria that the evaluators establish at the outset of the course and update periodically. Since each student's performance differs for the levels of achievement, evaluators map performance over all the levels and make trade-offs to assign the student a grade.[57]

Instructors in England, Scotland, and Wales implemented this assessment methodology in 1991, and researchers discovered several problems. Namely, instructors required additional support staff to help administer the SATs, and the assessments disrupted the normal classroom routines to the extent that classroom discipline deteriorated significantly. Further, the assessments required an inordinate amount of time to plan; to collect the necessary material; and to administer, grade, and record the results. Instructors had to concurrently teach students who were not taking the SATs and then restore the class to a normal schedule after the students completed the assessments. Finally, these assessments were expensive to develop and administer, and many instructors expressed dissatisfaction with the entire process.[58]

Evaluators also use criterion-referenced assessments in scientific or engineering disciplines. For example, they can examine the levels of achievement for an academic area called "practical science skills" to determine student performance in the laboratory. The achievement levels include: observing, recording, measuring, and manipulating and following procedures and instructions. Evaluators base their assessments of students on established performance standards.[59] They then can employ various techniques in science classes to assess student performance with respect to the achievement levels. In one such technique, evaluators use checklists to measure student performance against a number of predetermined criteria. Checklists are effective if evaluators assess well-defined and rather simple skills and abilities; however, evaluators find complex skills and abilities difficult to assess when the number of criteria becomes too large. Evaluators also can use structured or semistructured taskings as assessment tools by temporarily stopping the lesson and giving the students a short practical test item to complete. Finally, evaluators can measure the students' global performances by assessing how well they tackled an experiment or investigation. Unfortunately, research has shown this last technique to be unreliable and invalid.[60]

While criterion-referenced assessments compare information on student performance to specific standards, they have little predictive power since they only report achievements at a particular time. In addition, evaluators must precisely define the outcomes they assess. They must assess the accuracy of the students' answers, the amount of important and relevant information the students included in their answers, and the quality of their responses.[61] Finally, Grant Wiggins believes these assessments to be inadequate, as evaluators usually present contrived problems or situations with artificial cues. Authentic academic challenges are inherently ambiguous and open-ended; therefore, assessments should possess sophisticated criteria to determine how well students accomplish these challenges.[62]

Authentic Assessments

Assessments that are authentic are public and involve actual audiences, clients, and panels as the evaluators. Relying on predetermined standards and prior testing, these evaluators make judgments involving multiple criteria in assessing students. Authentic assessments are not needlessly intrusive, arbitrary, or contrived; they emphasize realistic complexities and stress depth more than breadth of knowledge. They are not structured for recall or lucky responses, nor are they fragmented or static tasks. Instead, these assessments resemble realistic situations and involve ambiguous tasks or problems intended to be contextualized and complex intellectual challenges. Finally, authentic assessments highlight the students' strengths by accommodating their learning styles, abilities, aptitudes, and strengths, thereby enabling them to demonstrate what they can do.[63] As Randy Elmore states,

> Conducting a science fair project, giving an art exhibition, making a speech, writing an essay or research paper, or demonstrating how to repair a small engine are examples of authentic [assessments]. Students must show real competence and not just an ability to recognize solutions from contrived questions. In authentic assessment[s], teachers examine the process of problem solving as well as answers selected.[64]

Application-oriented projects are authentic assessments which stimulate a student's creativity and can increase his or her motivation to learn. Evaluators also can use the projects to observe the students directly and assess their performance in, as Andrew Beale states, "Situations that demand application and transfer of knowledge, as well as noncognitive functions such as interest, attitude, cooperation, persistence, and the like."[65] Evaluators develop checklists to help them observe the students as they work on a project, thus reducing the potential for subjective distortions or bias. They can furnish these checklists to the students prior to the assessment so the students will know what is important, what will be expected of them, and what level of proficiency they will be required to achieve.[66] However, students may tend to prepare themselves to accomplish the specific items on the checklists, and the evaluators may not observe an unbiased or unprepared student performance.

Evaluators also can use authentic assessments to determine how business and office education students accomplish realistic tasks that are comparable to those performed in entry level positions in, for example, accounting, secretarial, and general office settings. Authentic language assessments include writing realistic letters to potential employers and conversing with trained interviewers. Evaluators can authentically assess science classes by tasking students to design realistic experiments, use equipment such as microscopes or construct such hardware as electronic circuits, machines, and structures.[67]

Designing authentic assessments requires innovative approaches in every discipline. For example, English instructors may wish to assess a student's reading skills and abilities by using realistic material or situations. However, while texts have topical and structural integrity, most reading assessments use short pieces of contrived texts that lack the integrity or cohesiveness found in real texts. In addition, even though reading orchestrates many skills and abilities, assessments fragment the reading of texts or passages into isolated skills to facilitate the process of assessing student performance. Although reading comprehension depends on prior knowledge, current assessments use short passages about unfamiliar topics to mask the effect of a student's level of knowledge on his or her ability to comprehend texts; therefore, they fail to assess the impact of that knowledge on his or her comprehension. Often, reading assessments provide a few sentences and require the student to ascertain a single thesis, make subtle inferences about concepts with which they are unfamiliar, and select answers that can deliberately mislead. Finally, inferential and critical reading are essential to construct meaning, but assessments predominantly use literal and sentence-level inferential comprehension items.[68] Valid assessments should "provide students with multiple opportunities to apply their reading skills to a variety of real-life texts and tasks."[69]

A portfolio system is one way evaluators can authentically assess a student's reading skills with a variety of tasks. A portfolio consists of multiple readings the students complete, including texts with various contexts and texts with different purposes or audiences. These readings incorporate various assessment measures that indicate such forms of expertise in reading as comprehension, uses of literacy, grammar, and metacognitive strategies. Students and evaluators frequently add readings to the portfolios so that the measures determine evolving knowledge and understanding.[70] According to Roger and Beverly Farr, "Students are encouraged to take responsibility for maintaining their portfolios, determining some of the contents, and evaluating their own progress."[71] Evaluators select narrative and expository texts which are representative of the types, content, and structure of materials students read in school, as they are more interesting and motivating for students and engage the students in more complex reasoning and thinking. With these more realistic samples of student readings, evaluators can construct inferential and critical reading questions which lead to authentic assessments of a student's reading skills and abilities.[72]

In using authentic texts as assessment tools, evaluators must realize that the text drives the focus of the questioning rather than the questions driving the construction or selection of the text. However, to truly grasp a given text, the students must not only understand the information that is explicitly stated in the text and integrate it into their comprehension of that text, but they also must use the information in ways that expand beyond the immediate text, such as formulating new ideas or concepts. Therefore, evaluators must develop questions that not only test the students' comprehension but also how they have used the text to enhance their knowledge and understanding. Consequently, evaluators need to determine the depth and breadth of a student's knowledge and understanding of passage-specific concepts before he or she reads the passage. The student must first answer questions about the important concepts underlying the passage's central themes before reading the passage. The evaluators then use this information to establish a base of knowledge and understanding to address student familiarity of the topic and to help evaluators determine a student's comprehension of the text and his or her ability to expand beyond that text.[73]

To help interpret a student's level of comprehension, the questions that accompany the selected text query the student about the text's organization, the purpose of charts and illustrations, and strategies the student can use to interpret an unknown or confusing phrase or situation. Other questions associated with the text ask students about their perceptions of the difficulty of the text and their reading behaviors related to the text. By directly relating these questions to the text, evaluators can measure how successfully students understand, comprehend, and manage reading strategies to meet the demands of various texts and tasks. They also can use this information to assess student literacy habits, attitudes, abilities, and self-perceptions with respect to reading performance.[74]

Authentically assessing a student's reading skills requires a thorough assessment methodology. Assessing student writing skills and abilities also necessitates in-depth assessments. Since an assessment of a student's writing determines the level of the student's writing skills and abilities, evaluators should examine samples of authentic student writing. They can develop writing exercises in the same manner as the reading assessments by using the portfolio system described above. When evaluators task students to write about topics related to their schoolwork instead of writing about contrived or artificial situations, the students will become more interested and involved in the assignment and will put forth a genuine effort to complete it.[75]

One way to insure student involvement requires that the students choose the specific writing samples for their portfolios based on criteria the instructors and evaluators establish. The instructors and evaluators then formulate a portfolio agenda, which allows them to assess the students' abilities as writers and independent thinkers and includes a plethora of products revealing the students' strengths and weaknesses in writing. The first part of this agenda can, for example, include an introduction where the

students describe themselves to the readers of the portfolios, identify their characteristic writing styles, and summarize the portfolio's contents. The portfolio then contains examples of timed first-draft writing, which requires students to formulate, organize, and express their thoughts within a specified time. Students can submit other samples of writing, including one which is a packet that demonstrates evidence of a complete writing process: prewriting, planning, writing, revising, editing, and rewriting. Another part of the portfolio can provide examples of students' beginning and end-of-course writings to show their behavioral and cognitive growth during the course. Other writing samples can include creative writings; these assignments allow for more spontaneity and creativity. Instructors and evaluators can include in the agenda the requirement for the students to select their best piece of writing along with a rationale for their choice. Finally, the instructors and students can work together to choose samples of writing for the students' portfolios, thus allowing the students to learn the value of collaboration in the writing process.[76]

A completed portfolio can provide evaluators with an accurate measure of a student's accomplishments and understanding of the writing process, as well as reinforce his or her learning process. Further, the portfolio can record the student's quantitative and qualitative performance over time and provide a framework for self-assessments of his or her writing styles, skills, and abilities. The portfolio requires the student to be actively involved in deciding which writing samples the instructor and evaluators should assess; it also permits continuous feedback between the instructor and student. This authentic assessment helps students to understand that achievement must be multidimensional and multipurpose to capture the complexity of writing or, for that matter, any other discipline.[77]

Objectively assessing such authentic assessments as portfolios is feasible as long as educators clearly define the assessment criteria. One methodology which develops common assessment criteria has evaluators form a blind panel to collectively assess student performances on a particular assignment; the authors of the assignment remain anonymous to the members of the panel. Initially, the panel receives samples of student papers to read and to practice grading; the results are shared with each other; and evaluators then calibrate each panel member's performance to insure consistent grading. The panel uses a numeric scale—say from 0 to 9—to grade each assignment; a "0" implies the students did not respond to the tasking, and a "1" through "9" score indicates increasing levels of development.[78] Multiple panel members grade each student's submittal, and they add the scores from the graders to provide a composite score. Evaluators can then assign grades based on the composite score. They can task the panel to assess the submittals based on such criteria as comprehension, style, grammar, mechanics, word choice, correctness and variety of sentence structures, organization, completeness, and level of thinking. When evaluators assess student competence in communication, they must establish a criterion level for adequate performance on each assigned task. Their definition of adequate performance

needs to account for such context features as whether the language in the submittal is appropriate to the tasking.[79] The evaluators must devote considerable time and effort into identifying the appropriate levels of performance, as the quality of the assessment depends on the thoroughness and accuracy of the performance levels.

Authentic assessments are not limited to reading or writing; mathematics instructors also use the portfolio as their authentic assessment methodology. The portfolio provides samples of how students solve math problems and how they communicate mathematics terminology, and it indicates the quality of the math course the students are enrolled in. Evaluators do not assess a student's performance on a numerical scale but against such specific criteria as the student's ability to understand and solve a problem; the student's observations, connections, and generalizations; terminology, notation, and symbols the student uses to communicate; clarity of his or her communication; and his or her use of graphs, charts, tables, models, and diagrams. Evaluators can assess the quality of the mathematics program by soliciting student comments and recording their responses in specific content areas, including estimation, measurements, relationships, operations, and theory.[80] Evaluators can use portfolios as effective tools to assess a student's level of achievement in mathematics or other science disciplines; however, the portfolios must be extensive, and they are quite time consuming to develop, administer, and analyze.

Performance Assessments

Closely related to authentic assessments are performance assessments. The National Council on Education Standards and Testing asserts that performance assessments

> require students to complete challenging tasks that call for deep understanding of subject matter, problem solving, and communication. These tasks may be conceived as extended projects, hands-on demonstrations such as conducting experiments, or portfolios, where students include evidence of a range of accomplishments or their developed expertise.[81]

Performance assessments ensure that students learn more than just basic knowledge and skills by providing them with the opportunity to use their knowledge, skills, abilities, and understanding to address new problems or situations. These assessments determine if students develop the ability, for example, to solve novel problems, work cooperatively in groups, or synthesize knowledge and comprehension across disciplines.[82] The assessment methodologies focus on stated curriculum goals and objectives, measuring both content and process and reflecting a high degree of fairness to all students; they differ from authentic assessments in that they do not constrain evaluators from using only authentic or realistic situations for the assessments. Performance assessments can use abstract concepts and scenarios to test students, and "encourage self-directed learning, . . . focusing on the process rather that the result of learning."[83] According to Jane Heckley

Kon and Giselle Martin-Kniep, "By offering a wider range of test formats, more students have an opportunity to show what they know and what they can do. This is particularly true for the assessment of higher-order thinking skills, for which the performance tests seem to be particularly reliable."[84] Courses in history and the social sciences use performance assessments to accomplish the following goals:

1. determine student learning needs
2. provide students with information and assistance on their progress toward curriculum goals and objectives
3. provide information for assigning grades
4. compare student achievement to an established norm.[85]

Evaluators of history and social science courses tailor tests to determine if students achieve the desired goals and objectives. They can use norm-referenced tests to assess student retention of knowledge and skills or they can use criterion-referenced tests to determine the extent to which students have mastered a specific number of behavioral and cognitive objectives. The evaluators augment these tests with such performance assessments as portfolios of student papers, oral discourses, projects, exhibitions, and essays which focus on higher levels of thinking. Often, students must produce a demonstration or live performance as a comprehensive and interdisciplinary demonstration of their skills, abilities, competence, initiative, and creativity. Thus, through performance assessments, evaluators expose students in history and social science courses to a comprehensive, systematic, and valid assessment process.[86]

Evaluators also use performance assessments in science classes. The ideal way to assess student performance in a science class would be to observe directly "a student pursuing a scientific inquiry with laboratory equipment and materials."[87] Unfortunately, such observations are costly, time consuming, and usually difficult to obtain. Hence, performance assessments combine various assessment methodologies to approximate the "ideal assessment." Such methodologies include

1. lab notebooks in which students record their observations and conclusions,
2. computer simulations of the students' hands-on investigations,
3. short answer paper-and-pencil problems dealing with planning, analyzing, and interpreting experiments, and
4. multiple-choice tests addressing specifics about conducting and analyzing hands-on investigations.[88]

Each of these methodologies possesses distinct attributes which contribute to make valid and thorough assessments of student performance. For example, evaluators can use notebooks as surrogates to actually observe students conducting an experiment, as long as the students explain in detail every step they take during the experiment. Evaluators can read the notebook to determine how well the students performed the experiment. The notebooks also provide the students with the opportunity to express themselves in

writing; writing is an important skill sometimes neglected in science classes. Computer simulations are less costly and time consuming than hands-on investigations, although the development costs for the simulations can be considerable. The computer also maintains complete records of student performance on specific exercises to aid in the evaluation process. Students have the opportunity to repeat expensive experiments on the computer or attempt alternative procedures which might not be possible with hands-on experiments. Paper-and-pencil problems and multiple-choice tests help to standardize assessments and are essential components of a complete assessment of the students' achievements and their levels of knowledge, skills, abilities, and comprehension. Evaluators must conduct these performance assessments repeatedly to reflect student achievement and progress accurately; single measures of performance cannot provide adequate or accurate information about student performance.[89]

Indeed, evaluators design performance assessments to provide an accurate picture of student performance. They also use them to assess a student's higher level of thinking and advanced cognitive skills and abilities. Many mathematics courses challenge a student's ability to think and reason by using problem-solving performance assessments. The problem-solving methodology allows a student to confront a novel situation, formulate connections between given facts, identify the goal, and explore possible strategies to reach the goal. One difficulty in assessing a student's ability to solve math problems results from the failure of the student to articulate what he or she is doing or thinking. In addition, evaluators tend to formulate general impressions of the quality of a solution while scanning the student's work; good solutions with minor errors—which dramatically alter the answer—may receive undeservedly low scores.[90]

Evaluators can assess a student's problem-solving skills and abilities accurately by using scales that focus on solution procedures rather than on the answer itself. These scales enable evaluators to establish fairer and more reliable scores; they assign separate scores to each of the three stages in the problem-solving process: understanding the problem, solving the problem, and answering the problem. Comprehensive descriptions of the performance levels for each stage provide the basis for the scales. Evaluators can expand the number of scales to assess the procedures students follow in formulating a solution to the problem; the evaluators can include such categories as strategy selection and strategy implementation. Instructors can focus on individual categories to determine how well their students understand specific aspects of the problem-solving process. When students realize that they are being assessed on the entire problem instead of the solution only, they become more motivated to communicate and explain their thinking.[91] Evaluators also can use this scale system to compare the effectiveness of different teaching styles. They can observe differences in student levels of understanding by comparing the scores of specific categories assessed for each teaching style.

Unfortunately, as promising as performance assessments are, they possess some disadvantages. For example, evaluators must devote tremendous time

and effort to the assessments to insure that they are unbiased, valid, reliable, fair, practical, and efficient. Often, students who successfully accomplish one assessment cannot do another. The format evaluators use to gather the assessment data influence student performance; one type of format or assessment may be favorable to some students while putting others at a disadvantage.[92] Finally, because performance assessments are usually more expensive and time consuming, they provide evaluators with fewer opportunities to assess student performance than traditional assessment methodologies.[93]

One contributor to the cost and time problem of performance assessments is the requirement for repeated, or even continuous, assessments to assess student performance accurately. Still, evaluators must continuously assess student performance to observe student growth in the subject matter and the course and to obtain a complete picture of the effectiveness of different teaching styles. Variations in the growth patterns of students in various learning environments may indicate differences in the effectiveness of different teaching styles.

Analyzing Assessment Data

Assuming that they can develop the optimum assessment methodology and then collect the appropriate data, educators may find it difficult to analyze the data to develop a student's growth patterns. According to Anthony Bryk and Stephen Raudenbush, evaluators use the hierarchical linear model (HLM) technique to conduct such an analysis. HLMs provide

> an integrated approach for studying the structure of individual growth, examining the reliability of instruments for measuring status and change, investigating correlates of status and change, and testing hypotheses about the effects of background variables and experimental interventions on individual growth.[94]

In other words, as a statistical tool, HLM uses multiple data points acquired with periodic assessments to track student growth in specific categories of performance. HLM also can establish common baselines to examine the effects of such diverse variables as student backgrounds or familiarity with teaching aids. Finally, evaluators can use HLM to discover factors that influence the rate at which students learn material and to assess the reliability of measures which examine student growth rates. Unfortunately, HLM results rely on the strength of the data. Inferences based directly on estimated variances and covariances are likely to be imprecise with small sample sizes. The size of the sample, as well as the accuracy of the data, influence the estimated reliability of the growth parameters.[95]

Qualitative Assessments

Even though HLM analyzes assessment data well, evaluators cannot apply it to assessment methodologies which do not generate quantifiable information.

Like the assessments previously described, qualitative methodologies assess student performance resulting from the students' exposure to certain styles of teaching. However, since these methodologies are qualitative in nature, they may not provide information which evaluators can analyze with such statistical tools as HLM. Still, they can use these methodologies to help form a complete assessment of student performance.

Focused Listing Assessment

For example, evaluators use focused listing assessments to determine student knowledge by asking the students to list ideas that are critically related to an important course topic. After the topic has been presented in lectures, discussions, or assigned readings, the instructor can task students to list important words or phrases they consider most important.

By comparing each student's list to a "master" list, evaluators can develop a clear idea of what information each student retained. Since focused listing indicates the most salient information from the student's viewpoint, it requires only recall and therefore makes no demands on higher level cognitive skills. Further, students may be able to produce a list of relevant terms without understanding their meanings or interrelationships.[96]

Directed Paraphrasing Assessment

Directed paraphrasing is similar to focused listening, but with this assessment methodology evaluators direct students to paraphrase a reading or lecture in their own words, directed to a specific audience, and constrained to specified page-length or speaking-time limits. Evaluators use the completed assignments to judge the degree to which students have understood and internalized a topic, and permit the instructor to determine quickly in some detail how much and how well the students comprehend the topic. It helps to develop the students' skill and ability to actively comprehend and communicate information they have learned, but it is difficult to establish quantitative and qualitative criteria to perform thorough and standardized assessments.[97]

Background Knowledge Probe

While evaluators use directed paraphrasing after an instructor has presented a topic, they rely on background knowledge probes to use at the beginning of a course or before the instructor introduces a specific topic in the course. For this assessment technique, the instructor prepares a set of simple, interrelated questions to provide feedback on each student's background knowledge before he or she begins teaching.

The probes not only provide rich data about the student's knowledge of the topic, they also can help the instructor to gauge the appropriate starting point and the level of teaching for the topic. By using the same or similar questions at the midpoint and the end of the course or topic, evaluators can assess changes in students' knowledge and understanding. Probes also indicate

student skill and ability in communicating what they know and encourage them to relate the topic to their own knowledge and experience.[98]

Documented Problem-Set Assessment

Background knowledge probes are intended to generate qualitative assessments of student knowledge and understanding and, as such, are difficult to use in such quantitative skills and abilities courses as algebra, calculus, and statistics. Evaluators can assess student performance in these courses by using documented problem-set assessments. For this methodology, the instructor tasks the students to solve several problems—within a specific time period—that are representative of the topic the students are learning. In the process of solving the problems, the students explicitly document the steps they took to arrive at a solution. By examining the responses, evaluators can discern both the students' ability to answer the given problems correctly and their approaches to problem solving. A documented problem-set assessment methodology focuses attention on each student's general skill and ability level rather than discrete, unique answers and allows both the students and the instructor to become aware of a multitude of successful, and unsuccessful, problem-solving techniques. However, students may find it difficult to document their work explicitly, and instructors and evaluators require a great deal of time and energy to prepare and assess the problem sets.[99]

Focused Dialectical Notes Assessment

Evaluators cannot use documented problem-set assessment methodologies with courses that require close examinations of texts; these courses include history, philosophy, political science, literature, and law. Evaluators can use the focused dialectical notes assessment in these courses, as this methodology tasks students to take notes on a selected reading passage. The students consider these notes as a conversation with the text; they record their reactions to the text, including agreements, disagreements, questions, and assertions. These dialogues can provide detailed feedback on how students analyze and respond to the academic texts they are reading. Evaluators can determine weaknesses in the students' analytic reading strategies in addition to the students' misunderstandings of the content of the text. While the notes provide to evaluators contextualized samples of students' skills and abilities at reasoning and analyzing the readings, students may try to write what they think the instructor wants rather than writing their candid responses.[100]

Invented Dialog Assessment

Another assessment methodology that evaluators find useful in humanities and social sciences courses is the invented dialog. Here the instructor selects one or more controversial issues, theories, decisions, or personalities related to the topic and requires the students to develop original, lively, and persuasive dialogues about them. Students can speculate on possible but unrecorded conversations relevant to a given situation or juxtapose times and

places to change history and achieve their efficacy. They may select and integrate quotes from primary sources or invent reasonable quotes that fit the character of the speakers and the context.

Invented dialog draws on higher order thinking abilities, forcing students to internalize and comprehend course materials in different ways. Evaluators can assess the students' abilities to assume the intellectual personalities and styles of expression of other people and the students' understanding of theories, controversies, and opinions of those people. The students must present the material they have studied creatively, allowing the evaluators to examine their creative skills and abilities. Unfortunately, students who feel they are not creative or cannot write may resist this technique.[101]

Concept Map Assessment

Students who are uncomfortable with writing may appreciate concept map assessments. According to K. Patricia Cross and Thomas Angelo, "Concept maps are diagrams that students draw in response to a stimulus word or phrase. The 'maps' illustrate the associations students make between the stimulus and other words or phrases—the latter generated by the students themselves."[102]

This methodology is particularly appropriate for courses with high theoretical content. Evaluators can employ concept maps before instructors present a topic to discover the students' preconceptions and prior knowledge and comprehension. In using the maps before, during, and after the topic is presented, evaluators can assess changes in students' conceptual perceptions. They use concept maps to record the patterns of associations students make about a given focal concept. Evaluators also can use them to assess the students' understanding of relevant conceptual relations and how much that understanding corresponds to the instructor's or the discipline's model. The maps provide a graphic view of students' conceptual views and force them to consider how their ideas and concepts of the topic are related; however, in-depth examination of this methodology may require a sophisticated statistical analysis.[103]

Student and Faculty Evaluations

Ideally, the assessments described above should match the teaching styles the students are exposed to and include hands-on performance tasks. They should probe a student's depth of understanding and ability to master a body of knowledge. In addition to judging the correctness of an answer or a performance, assessments should emphasize how the student ascertained the answer or how the student carried out the activity.[104] However, assessing student performance in a learning environment is only one facet of thoroughly assessing the effectiveness of different teaching styles. Student and faculty evaluations also allow evaluators to assess styles of teaching in institutions of

higher education.[105] Most current evaluations examine student and faculty opinions of instructor performance and effectiveness in the classroom. Although evaluations associated with this research project must examine the effectiveness of the teaching style itself, educators can consider most evaluations of instructors to be equivalent to evaluations of teaching styles and modifiable where appropriate.

Note that the performance judgements of instructors that are made by students, faculty observers, and the instructors themselves vary considerably. Self-evaluations are the most difficult means of evaluating instructor performance, and they require from the instructor a high degree of maturity and objectivity.[106] Instructors have erroneous perceptions of how others perceive and assess their performance; the ratings they give themselves are higher than the ratings given to them by their students or faculty observers. Colleague evaluations based on classroom observations are generally not reliable due to the usually small number of visits by faculty observers and their relative inexperience in evaluating teaching performance. However, faculty members can evaluate instructor performance effectively if they are trained evaluators and if evaluators increase classroom visitations, both in frequency and in the number of faculty members who visit the classroom.[107]

Reliable colleague evaluations substantially mirror student evaluations of teaching performance; however, evaluators usually use student ratings more frequently to evaluate instructor performance. Herbert Marsh and Michael Bailey state that "instructors appear to have distinct profiles of strengths and weaknesses that are highly generalizable and that students are apparently able to discriminate their instructors' strengths and weaknesses, at least when ratings are aggregated over many students."[108] Further, they have concluded that students' evaluations of teaching effectiveness are:

1. multidimensional,
2. reliable and stable,
3. primarily a function of the instructor who teaches a course rather than of the course he or she teaches,
4. relatively valid against a variety of effective teaching indicators,
5. relatively unaffected by potential biases to the students' ratings, and
6. perceived to be useful by faculty as feedback tools.[109]

While the consensus opinion of evaluators is that students can rate the performance of an instructor effectively, caveats do exist. To enhance consistency in student ratings, evaluators should administer and collect appropriate forms in a prescribed and systematic manner. The students should complete these forms several times during the course; most students do not take end-of-course evaluations seriously, since the results will not effect them. Research indicates that such variables identified on evaluation forms as ability level, sex, or class level do not produce significant differences in the ratings. However, small classes usually generate especially high ratings, and major courses tend to receive higher ratings than core courses; these trends are logical because students are more involved and motivated

when compared to students enrolled in core courses or large classes. The image students hold of an ideal instructor influences their ratings. Like colleague evaluations, students are not trained evaluators; hence, some educators question how well students can actually rate instructor effectiveness. In many cases, students' ratings tend to correlate with their achievements in class. Still, research supports the use of both reliable faculty observer and student evaluations to assess instructor performance.[110]

Types of Evaluation Forms

Evaluators have developed several different types of evaluation forms to measure the performance of an instructor. These forms range from checklists and standardized forms to questionnaires and rating guidelines. For example, the check scales form lists instructor attributes—as well as functions and outcomes of the course—on a standardized form and tasks the evaluator to rate each item separately. A characterization report differs from the check scales form in that it rates the total merit of the instructor according to a scale of values. The guided comment report lists a series of topics or questions for the rater to use in writing about the instructor. The descriptive report is similar to the guided comment report, but is less structured; in fact, a descriptive report does not specify topics and leaves the structure of the statements to the rater's discretion.[111]

A contentious issue with these evaluation schemes concerns the criterion against which a rater makes his or her evaluation. What type of instructor or style of teaching does the rater hold as a model? The rater's attitudes and opinions can bias his or her evaluation greatly; untrained people who assess others compound the possibility of this bias. A halo effect, where they see all individuals in a particular light, influences some raters. Some raters tend to rate an instructor as other raters do, perpetuating that instructor's perceived success or failure in teaching. While objective evaluation forms may counter some bias, the evaluators who examine the completed evaluations must look also to see what viewpoint the rater used in the evaluation.[112] For example, students perceive teaching differently; they respond in various ways to techniques instructors use to stimulate interest, to discussion, and to the classroom atmosphere. These individual differences influence their ratings of the instructor's effectiveness.

The rating scales and wording on the evaluation forms also influence the rater. Scales often fail to define the entire scope of an instructor's teaching efforts; therefore, the raters cannot use the forms to fully evaluate teaching performance.[113] Hence, evaluators must judiciously choose the appropriate scales and wording to provide the best possible evaluation forms for the raters.

Evaluators use three widely distributed response formats for evaluations: verbal scales, end-anchored scales, and numerical scales. The verbal scales format gives a simple verbal label for each scale point; it uses no numbers. A variation of verbal scales, the end-anchored scales format gives labels to the end points of the scale but has no labels for any of the points in between.

However, evaluators might assign numbers to the points on the scale. Similarly, the numerical scales format numbers each point on the scale but does not verbally describe any of the points; the instructions on the evaluation form may specify which point is lowest on the scale.[114]

Research shows no difference in either the verbal or end-anchored scales in terms of overall mean or degree of accuracy of the evaluations. On the other hand, with numerical scales, the overall mean and degree of accuracy are higher when compared to the first two scales. In fact, researchers have found it desirable to use rating scales where each point has precisely defined numbers and labels.[115] For example, the six-point response scale the Educational Assessment Center at the University of Washington developed is: excellent, 6; very good, 5; good, 4; fair, 3; poor, 2; and very poor, 1.[116]

Still, numerical and labeled ratings have different meanings to different people, and evaluators can misinterpret the rater's reactions if they base the evaluations on the ratings alone. To encourage complete and detailed feedback, evaluators must provide additional space on the evaluation forms for a rater's comments. The raters can verbalize their rationale for their ratings. By including additional questions on a wide range of topics that scales cannot address adequately, raters can comment and make constructive criticisms to provide considerable detail in their evaluations. They must have ample space on the forms for their comments, as raters tend to limit their comments to the available space. When students evaluate instructors, evaluators should give them enough time to fill out the evaluation forms, and the evaluators should emphasize the importance of the task. Students will then tend to spend more time completing the forms and to offer more complete and detailed feedback.[117]

Types of Questions for Evaluation Forms

Student and colleague feedback, however, is only as complete as the evaluation form itself; therefore, choosing the proper questions is critical to the evaluation process. Many examples of evaluation questions exist and have been used in diverse situations. Evaluators must select those questions they feel will determine student and faculty observer opinions of teaching performance and effectiveness. For example, one end-of-course student survey consists of seven statements designed to evaluate course and instructor characteristics and two questions designed to provide the student's overall evaluations of the instructor and the course. Table 1 shows the statements and questions presented to the students. Students respond to the statements by using a four-point Likert-type scale, and they answer the two questions with a five-point Likert-type scale.[118] The number and labels associated with the five-point scale could range from 1 (improved a great deal) to 5 (showed no improvement at all), or from 1 (little) to 5 (a great deal).[119] The four-point scale would resemble the five-point scale; the choice of numbering system and labels depends on the evaluators' preferences.

Evaluators use the same Likert-type response scales for the Teaching Analysis by Students (TABS) questionnaire developed at the University of Massachusetts' Clinic to Improve University Teaching. This extensively used questionnaire consists of 20 items that specify teaching behavior and four overall ratings of teaching. The labels for the four-point scale used with the 20 items include: no improvement needed; little improvement needed; improvement needed; and considerable improvement needed. Evaluators use a five-point response scale for the overall ratings to compare the course to other courses the students have taken. Table 2 summarizes the TABS questionnaire.[120]

Table 1

End-of-Course Survey Statements and Questions

COURSE AND INSTRUCTOR CHARACTERISTICS	OVERALL ASSESSMENTS
The instructor presented ideas and theories very clearly.	Overall, how would you rate the instructor?
The instructor was open to other viewpoints.	What is your overall rating of this course?
Classes were profitable and worth attending.	
I would rate the subject matter of this course as very interesting.	
I was satisfied with the amount of interaction I had with the instructor during the course.	
I had ample opportunity to ask questions during classtime.	
Students were strongly encouraged to think for themselves.	

The Student Evaluations of Educational Quality (SEEQ) instrument is another student evaluation form. SEEQ's 35 items measure nine evaluation factors. Table 3 lists these factors and items. Evaluators have used the SEEQ at different teaching levels, as well as in various academic disciplines, and more than 30 exploratory factor analyses have supported its implementation. While evaluators use this tool to evaluate an instructor based on nine different factors in a multidimensional construct, other evaluation tools they use assume that students' evaluations of teaching effectiveness can be viewed from an almost unidimensional perspective.[121]

For example, the 21-item Endeavor evaluation instrument defines seven distinct components of students' evaluations of teaching effectiveness: student accomplishments, organization/planning, presentation clarity, personal attention, grading/exams, class discussion, and work load. However, evaluators group these components into two global factors—pedagogical skill and rapport—which supposedly describe effective teaching. Similarly, the

Feldman evaluation instrument groups the 20 characteristics of a superior university teacher, shown in table 4, into three categories related to the instructor's role as a presenter, facilitator, and manager or director. The argument for narrowing a group of evaluative ratings into a single or small number of categories is that decreasing the number of categories eases personnel decisions for administrators. Unfortunately, evaluators lose considerable information when they summarize students' evaluations of teaching effectiveness in such a manner. Research indicates that a multidimensional perspective is more appropriate; evaluators require a variety of categories to gain a complete view of student impressions of an instructor's effectiveness.[122]

Table 2

TABS Questionnaire Items

TEACHING BEHAVIOR	TEACHING BEHAVIOR	OVERALL RATINGS
inspiring interest	atmosphere	how much learned
interest introducing topics	creativity	how significant is content
action when students bored	participation in discussion	effectiveness of teaching
challenge point of view	clarifying relationships	overall value
adjusting rate	clarifying material	
distinction between topics	explaining what is expected	
wrapping things up	clear purpose of class	
use of class time	evaluation consistent	
how performance is evaluated	keeping informed of progress	
exploring viewpoints	selecting material	

Specifically, ratings from multidimensional student evaluations like the SEEQ are more useful feedback tools than evaluations which provide a small number of global ratings or a single total score rating. While the global ratings may indicate the students' perceptions of an instructor's overall effectiveness, they have little diagnostic value. In other words, the ratings cannot identify specific areas of an instructor's teaching which needs improvement or where the instructor's strengths lie. In contrast, evaluators can use multidimensional factors to build rating profiles which emphasize the instructor's strengths and weaknesses in relation to different factors. Since the evaluators base these profiles on ratings averaged across different classes, they can use them to compare against future ratings as a means to track the instructor's progress.[123]

Table 3

SEEQ Evaluation Factors and Items

LEARNING/VALUE	INSTRUCTOR ENTHUSIASM	ORGANIZATION AND CLARITY
course stimulating and challenging	enthusiastic about course	lecturer explanation clear
learned something valuable	dynamic and energetic	materials well explained/prepared
increase subject interest	enhanced presentation	course objectives stated/pursued
learned/understood subject matter	teaching style held your interest	lectures aided note-taking
overall course rating	overall instructor rating	
GROUP INTERACTION	**INDIVIDUAL RAPPORT**	**BREADTH OF COVERAGE**
encourage class discussion	friendly towards individual student	contrasted various implications
students shared knowledge/ideas	welcomed students seeking help/advice	gave background of ideas/concepts
encouraged questions/answers	interested in individual student	gave different points of view
encourage expressing ideas	accessible to individual student	discussed current developments
EXAMS/GRADING	**ASSIGNMENTS/READING**	**WORKLOAD/DIFFICULTY**
exam feedback valuable	readings/texts were valuable	course difficulty (easy-hard)
evaluation methods fair/appropriate	they contributed to understanding	course work load (light-heavy)
tested course content as promised		course pace (slow-fast)
		hours per week outside of class

Table 4

Feldman's Categories of an Effective Instructor

PRESENTATION	FACILITATION	REGULATION
stimulates interest	sensitive to class progress	clarity of objectives
enthusiasm	class discussion	fairness, impartiality
subject knowledge	intellectual challenge	value of course materials
intellectual expansiveness	respect for students	supplementary materials
preparation and organization	availability, helpfulness	classroom management
clarity and understandable		feedback to students
elocutionary skills		difficulty and work load
sensitive to class progress		
perceived outcome/impact		

Analyzing Evaluation Forms

Regardless of the categories used, analyzing the evaluation forms plays a crucial role in the process of assessing the effectiveness of different teaching styles. Evaluators can use ANOVA and ANCOVA to analyze the data from the forms and present the results in an appropriate manner for decision makers and information users.[124] Evaluators also use confirmatory factor analysis (CFA) and hierarchical CFA to statistically analyze the evaluation forms. These tools compare the ability of existing models to fit the observed data; therefore, evaluators can manipulate them to compare teaching effectiveness.[125]

While each statistical technique has inherent advantages and disadvantages, evaluators must ensure that they use the appropriate technique for the data they have gathered. Further, considering each student as a unit of analysis produces significantly different results than using a group of students as the unit of analysis. According to Patricia Cranton and Ronald Smith,

> When individual students within one class are the units of analysis, the variation in ratings reflects individual differences in the perceptions of students. When class means are the units of analysis, the variation should reflect perceived differences among instructors. Deviations of ratings from class means should reflect individual differences. . . . When individual student ratings across different classes are used, the variation due to instructors cannot be separated from the variation due to individual perceptions.[126]

Some researchers believe that the nature of the items in the evaluation forms influence the students in their ratings; hence, they question the validity of the forms themselves. Student interest and involvement, classroom atmosphere, class size, level of teaching, and discipline also are related to the way students perceive teaching, and, hence, how they evaluate instructors. Using the class mean score as the unit of analysis, however, diminishes these effects, since it eliminates individual differences among students' perceptions.[127]

As previously mentioned, objective evaluation forms may eliminate individual differences in students' perceptions. Unfortunately, these rating forms are usually standardized, and they do not provide the feedback that students prefer. Although many institutions use standardized forms, researchers have discovered that students feel that their evaluations do not directly impact the instructors or the courses, and, therefore, do not put much effort in completing them.

Alternatives to Evaluation Forms

Alternative methods for collecting student opinions directly involve students in the evaluation process. For example, evaluators can use an interview as a face-to-face survey or questionnaire with a more flexible approach to soliciting student opinions of different teaching styles. They may choose to interview students with diametrically opposed points of view, and the substance of the

interview may vary between respondents. Interviews can more closely examine opinions and attitudes towards the teaching styles or learning environments than surveys and questionnaires.

Interviews

A major advantage of the interview is the evaluator's capability to follow up on specific responses or elicit more complete responses.[128] In using interviews, Linda Lederman states that evaluators must

> presume that people (a) are able to recall and articulate their perceptions and feelings; (b) that they have the desire to be honest. . . . [However,] people may or may not be willing and or able to report with accuracy upon their own thoughts, feelings, and behaviors; and . . . there may be a discrepancy between what people believe they do and their actual behaviors.[129]

Focus Group Interview

Evaluators use the Focus Group Interview (FGI) technique to gather qualitative data on teaching effectiveness. They use in-depth, group interviews rather than individual interviews to elicit opinions from a deliberate but not necessarily representative sampling of students in a class. The group setting can provide a safe atmosphere for the students, and they can consequently interact with each other, as well as with the interviewer; the synergy can generate more information than the sum of individual interviews. Hence, the generated data can be deeper and richer than that elicited in individual interviews, and students may find it easier to express their thoughts, feelings, and experiences when they are in the presence of other students.[130]

In addition to eliciting more information from students, FGIs permit evaluators to gather more data in a shorter time than is possible in one-on-one interviews; with individual interviews the evaluators must address each topic with each student rather than once with the group. Evaluators can also observe the students' interactions with each other, thus generating additional data not available with individual interviews.[131] Conversely, FGIs can negatively impact the collection of student opinions. Some effects include "the presence of others on the individual participant's ability to express himself/herself; the extent to which the presence of others have unintended or unwanted effects on individual behavior; and the extent to which other group-related dynamics may intervene in any specific data collection."[132] The interviewer can, however, control these adverse effects by interceding if all members of the group are not participating equally.

Evaluators focus the FGI on specific topics they must examine, and establish homogenous groups for the interview. They recruit a limited number of students from the class they are examining based on a screening process which insures that group members meet a set of predetermined criteria. Evaluators determine the size of the group based on such practical limits as number of available students, cost, and time. The interviewer uses a guide so that he or she can cover the appropriate topics in enough depth to generate

the desired information. The guide is not a rigid and restrictive agenda but frames a series of questions to elicit student responses. The interviewer can conduct the FGI in formats ranging from structured and directive to semistructured and nondirective, depending on the questions he or she asks and the nature of the data the evaluators desire. The FGI produces qualitative data which evaluators use to generate an interpretation of or insight into specific aspects of student thoughts, feelings, and opinions relating to teaching effectiveness. The students respond to issues and materials emotionally and intellectually and can express themselves in an open environment.[133]

Small Group Instructional Diagnosis

A variation of the FGI, the Small Group Instructional Diagnosis (SGID) allows a facilitator to work directly with the instructor and the students in the class. The facilitator and instructor determine how they can best use a class interview to provide feedback to the instructor. With the instructor absent, the facilitator interviews the class members to obtain their opinions and views of the instructor and the style of teaching. The students form small groups to develop consensus on the strengths of the class, areas for change, and recommendations to the instructor for making the changes. The facilitator summarizes the groups' ideas, clarifying where necessary, so the groups are satisfied that the facilitator understands their opinions. The facilitator and instructor then meet to discuss the results and to develop a teaching improvement process that is responsive to the students.[134]

The SGID process differs from evaluations which use standardized rating forms in that a facilitator interviews groups of students to identify relevant teaching issues; the instructor and the facilitator usually conduct the process at the midterm of the course so that the instructor can react to the students' opinions while they are still enrolled in the course.[135] Students prefer the SGID process "because of the timing, quality of feedback, oral exchange of information and personal approach involved."[136] Research has shown the individual, standardized rating form evaluators administer at the end of the course as the least satisfying evaluation process mainly because of the instructor's limited reaction to the information the students provide. Students prefer to voice their opinions at midterm and then receive extended reactions from the instructors. Researchers also have discovered that students prefer direct feedback from the instructor, and they want to know what other students tell the instructor about the course. Class interviews permit students to generate their own response categories and to check their perceptions of teaching effectiveness with other students; hence, students prefer interviews as a means of collecting opinions about the course. If evaluators use standardized student rating forms, they should distribute them at midterm, and the instructor should provide extended reactions to the students' opinions.[137]

While students may favor the interview process to express their opinions and attitudes openly, evaluators may find interviews costly and time-consuming. Evaluators must be trained interviewers who can keep the interview on track and elicit beneficial and productive responses from the students. Interviews invariably take a long time for evaluators to organize, conduct, and analyze. They must develop clear and nonprejudicial wordings for the questions they ask, and they must pay attention to their demeanor and manner and to the students' nonverbal responses. Evaluators may find it difficult to interpret the students' responses and organize them in a concise and logical format. Finally, they must verify the consistency of their results across different interviews and settings.[138]

Evaluation Miniforms

Evaluators who prefer written feedback from students as opposed to oral interviews still can avoid the standardized rating forms by using instructor-designed evaluation miniforms. They can use these forms to collect limited, focused data on students' reactions to questions instructors consider important. The instructor develops from three to five specific questions for students to answer in a desired format, including multiple choice or scaling and administers the questionnaire at regular intervals to measure progress and attitudinal changes.

Evaluators format the questionnaire so that it easily fits the needs of the instructor and the class. They can devote their attention to the direction, intensity, and consistency of responses rather than to the actual ratings on the questionnaires, perhaps observing patterns in the student opinions.

If the instructor carefully constructs the questions, the students provide context-specific responses which result in focused feedback. However, generating such questions requires a great deal of thought, and if instructors fail to adequately develop them, students may not respond constructively to the intent of the questions.[139]

Electronic Mail

Although the miniforms provide individualized feedback to the instructor, evaluators may find them difficult to use to draw useful conclusions about how instructors can improve their teaching or their classroom setting. The interview process can generate this information; evaluators also can use electronic mail to elicit and evaluate students' reactions to instructors.

Electronic mail allows the evaluators to raise questions about classroom teaching and encourages students to respond to the questions by writing anonymous open letters, through the school's electronic mail system, to the evaluators. The letter format allows for diversity in content, tone, and length, and encourages direct, personal, and candid responses. However, the responses may be more critical and less beneficial than the evaluators desire. Students also may harbor inhibitions about expressing themselves if they feel that their writing styles reveal their identities.[140]

One-Minute Paper

A variation of using the electronic mail system to collect written student feedback is the one-minute paper. At the end of a particular class, the instructor poses one or two questions to which the students are requested to respond in one minute. The questions can address the teaching style or such specific elements of the course as classroom procedure, course content, materials, activities, and assignments.

The instructor receives immediate feedback which evaluators can read and analyze in a short period of time. One-minute papers encourage students to become involved in the class, but questions may elicit undesired feedback. Students may be evasive if they feel the instructor can recognize their writing style or handwriting. Finally, posing the appropriate but quickly answered questions offers a significant challenge.[141]

Chain Note

Another technique evaluators use to elicit open student responses is the chain note. In this technique, the instructor writes a question about the class on a piece of paper, places it in a large envelope, and passes the envelope to the students. The students respond to the question on index cards and place the cards in the envelope. The chain note elicits limited information about what each student thinks or notices about the style of teaching or the instructor at a particular instance during the class. It also lets the evaluators know what holds the students' attention during class by acting as a gauge of each student's level of enjoyment and involvement.

Class notes encourage student reflection, evaluation, and active observation and promote spontaneous and honest responses from the students. Students usually offer more concrete and specific reactions than with end-of-course evaluations, and evaluators can note positive and negative trends in the class. Unfortunately, class notes can distract the students from the information the instructor is presenting. Students may not respond if they believe the instructor can recognize their handwriting, and evaluators may not be able to interpret the notes if they are fragmented.[142]

Assessing Teaching Effectiveness Based on Program Cost

Soliciting student and colleague feedback is indeed one way to assess the effectiveness of an instructor or a style of teaching; assessing student performance is another way. However, in many situations, the measure of effectiveness is economic efficiency; rather, is the style of teaching or learning environment worth the investment in time and money? Evaluators cannot express a teaching style in terms of dollars and cents. They can, however, establish the costs of the instructors and the physical attributes of the

learning environment. While instructor or learning environment costs are convenient ways to determine if the decision makers should continue a particular education program, evaluators cannot assign a monetary value to quality. Similarly, it is difficult, if not impossible, to quantify the benefits of the teaching style, instruction, or learning environment in economic terms.[143] Finally, according to Nancy and Gary Padak, "When evaluations are tied to requests for funding, natural tendencies to highlight the positive and downplay the negative may overstate successes and obscure the program's weaker aspects."[144]

Still, many administrators and educators insist on assessing the effectiveness of teaching styles in terms of cost trade-offs. Given the current outlays for education, together with the current pressures for cost controls, an inevitable question seeks to determine what the academic institution receives for its investments in time and money. Evaluators can use several tools to help answer this question.

Cost-Benefit Analysis

Cost-benefit (or utility) analysis is a powerful tool for expressing the outcomes of personnel programs in monetary terms. This linear regression-based decision theory model uses a capital budgeting framework to assess the investment in an educational program. The acceptance criterion evaluators apply to the educational program is that the net present value of the program must be greater than zero. This net present value includes the program cost, the benefits derived during the assessment period, the duration of those benefits, and the discount rate representing the institution's minimum expected return on investment. Any benefits from the program ultimately must be stated in terms of direct, measurable changes in the institution's cash flow. The discount rate can be expressed according to the risk associated with those benefits that the investment will generate. A break-even analysis determines the minimum improvement in performance that yields the minimum required benefit. This minimum performance improvement can simply be the difference in performance between the experimental and control groups. Evaluators then can express the gains from an effective educational program in terms of two metrics: dollars and percent increases in student performances.[145]

Consensus Accounting Model

Cost-benefit analysis assumes that educational programs are investments, and, as such, evaluators should assess the programs in terms of economic costs and benefits.[146] Unfortunately, they may find it difficult to express education in terms of returns on investment. The consensus accounting model establishes an accounting measure for an educational program and does not attach a monetary value to performance. This model calculates the total cost for an educational program to make a direct comparison between competing programs based only on costs. It also assesses a program by monitoring its costs.[147]

Identifying the costs associated with an educational program is the first step in developing a consensus accounting model. Evaluators can identify two categories of cost: direct and indirect. Direct costs include those associated with the personnel involved in the educational program, outside goods and services, facilities, travel, per diems, accommodations, and incidental expenses. Evaluators group in the direct personnel costs the salaries and employee benefits of those people directly involved in the program. They also must include the salaries and benefits of personnel who assist the direct staff. Outside goods and services costs include expenses for designing, developing, reproducing, distributing, and reviewing educational materials and supplies for the particular program. Evaluators consider equipment purchased or rented for the specific educational program as a direct cost. They cannot trace indirect costs directly to a specific educational program. These costs include overhead, facilities costs, general and administrative expenses, and miscellaneous costs. Overhead usually refers to such physical items as the educational institution's share of organizational materials, equipment, and facilities. General and administrative expenses relate to the salaries and overhead for administrators, secretaries, technicians, and other people indirectly involved with the educational program. Once evaluators categorize the various direct and indirect costs, they can calculate and tabulate them to track individual costs and compare specific costs of different programs.[148]

While the consensus-accounting model provides a means to compare the costs of different educational programs, it does not assess the effectiveness of teaching styles with respect to cost. Rather, it embodies a tool to inform administrators of how much money a particular program costs. Again, it is difficult to assign a monetary value to a quality such as an education. Even so, cost is another element to consider when comparing different educational programs or learning environments.

Assessing the Material for a Learning Environment

A key factor in an effective learning environment is the quality and type of material the instructor and students use; the choices and use of appropriate materials are often crucial to the success of an educational program. Evelyn Nunes believes that

> There needs to be a more conscious attempt to make materials evaluation a planned, systematic, and ongoing part of an educational program. Determining which materials are suitable and effective is a concomitant of program objectives, goals, teaching styles, learning styles, and the needs of the target population.[149]

The process of choosing materials for a course is usually intensive and iterative. Students and instructors possess unique personalities and varying skill and ability levels which impact on the effectiveness of the material they use in a class. The materials also possess distinguishing characteristics which allow instructors to gear them toward those student characteristics, needs,

and interests which match their style of teaching. However, educators must determine if the materials instructors use are the most effective choices for the particular learning environment; therefore, they must develop tools to evaluate the effectiveness of the material.[150]

One straightforward evaluation methodology uses questionnaires and surveys to poll instructors and students on their impressions of the material they use in a particular class. The questions evaluators pose can resemble those they use to evaluate instructor effectiveness but are geared to evaluate such criteria as:

1. enjoyment and stimulating activities
2. appeal
3. extensive and appropriate selection of topics
4. relevance and objectivity
5. format is organized and in a logical progression
6. easy-to-follow directions
7. instructor and student can adjust difficulty level
8. degree of interaction between instructor and student
9. print and illustrations large and detailed enough
10. screen or text readable and uncluttered
11. sound and graphics enhance or detract from the lesson
12. instructor and student can control rate and sequence of program
13. material within the confines of the budget.[151]

These criteria address not only print materials but also the equipment, hardware, and software associated with educational technologies. Evaluators must consider all types of materials instructors use in a classroom since they contribute to the success or failure of the class. Materials for educational programs are investments; as such, administrators and educators expect the materials instructors use to yield positive results. Therefore, material evaluation must be a systematic and continuous element of a thorough methodology to assess the effectiveness of different teaching styles.

Assessing a Learning Environment's Physical Attributes

A final aspect to consider in assessing the effectiveness of different teaching styles is that of the learning environment's physical attributes. Dan Surry, a graduate assistant at the Air Force Academy, recently developed a plan to assess, in a technology-oriented basis, the networked classroom system (NCS) at the Air Force Academy. His plan examines such factors as how much the students and instructors enjoy the classroom, how comfortable it is, and how effectively instructors utilize it. The Academy wants to compare the NCS to traditional classrooms to see how this state-of-the-art, networked, multimedia classroom impacts student learning and instructor teaching.[152]

The NCS assessment plan consists of several assessment methodologies. The first one includes the instructor implementation log. Each instructor using the NCS completes a log that explains his or her daily classroom use. The logs' entries address equipment problems, comfort level, useful seating configurations, effective teaching styles, uses of the available technologies, and innovative classroom activities. The instructor also can identify ideally suited topics for the NCS and those which could be taught better in a traditional classroom. Instructors can use the data collected from the logs to modify and improve the courses they teach in the NCS and to develop other courses which they could teach in the NCS. The logbooks also provide a record of how the class progressed. Finally, evaluators and researchers can use them to guide changes and modifications to the NCS.[153]

Evaluators also use the one-minute paper to assess the NCS. The instructor uses the one-minute paper periodically to determine what the students felt was the most important lesson they learned during a specific class. Evaluators can use the information collected from students to determine if the students learned what the instructor intended for them to learn. Differences in the perceptions of the students and the instructor possibly can result from the influences of the NCS. In addition to the one-minute papers, students complete questionnaires three times during the course. These questionnaires determine student attitudes about the NCS itself. Similar to the instructor logbooks but written from the students' perspective, the questions relate to such physical conditions as temperature and seating comfort, such technical concerns as whether the equipment is user-friendly and if the projection screens are viewable, and such affective outcomes as student enjoyment level and motivation level resulting from using the classroom.[154] The questionnaire asks the students the following:

1. Do you think you are learning more in this course than you would if the class was held in a traditional classroom? Why or why not?

2. Are you enjoying this course more than you would if the class was held in a traditional classroom? Why or why not?

3. In your opinion, would it be a good idea for the Academy to install another networked classroom system? Why or why not?

4. What's one factor about the networked classroom that you especially like?

5. What's one factor about the networked classroom that you especially don't like or understand?

6. Is there any equipment or software in the classroom that doesn't work or works incorrectly?

7. Can you think of any equipment or software that could be added to improve the classroom?

8. What are your overall impressions of the networked classroom?

9. Please tell us about the physical environment of the classroom. (Are the chairs comfortable, is the room too hot or too cold, are there any distracting noises, etc.)[155]

The data evaluators compile from the questionnaires should reveal common complaints and provide them with an idea of the students' overall reactions to the classroom. In addition, the evaluators can use the questionnaires to determine changes which, from the students' perspective, educators need to make to the NCS. However, since the evaluators administer the questionnaires only to students who use the NCS, they have no comparative control group. On the other hand, the NCS assessment plan uses end-of-course surveys to compare student opinions with those of students enrolled in identical courses instructors teach in traditional classrooms. The Air Force Academy has used these standardized rating forms for some time, and, consequently, there exists a tremendous data base to compare with the opinions of students who are exposed to the NCS. Evaluators can possibly use the information collected from these surveys to determine if the students perceive any positive affective outcomes from the NCS.[156]

A fourth assessment methodology implemented in the NCS assessment plan is the instructor focus group. At the end of each semester, the instructors who used the NCS meet as a group to discuss their experiences with the networked classroom. They relate their day-to-day use of the classroom and discuss equipment needs, problems, suggested improvements, physical room arrangements, time considerations, effective teaching styles, and judgments as to the overall worth of the NCS. Evaluators and researchers can use the information they gather from this meeting to make changes to the NCS and to provide useful information to instructors who will use the NCS in subsequent semesters. The instructors also complete a questionnaire designed to elicit specific comments about the NCS. The questions include the following:

1. What was the most common way that you used the networked classroom; to give lectures, conduct demonstrations, assign student projects, etc.?

2. Do you think the students enjoyed having class in the NCS more than they would have in a regular classroom?

3. Do you think the students learned more because classes were held in the networked classroom?

4. Were the equipment and software in the classroom integrated and compatible?

5. How much material did you cover in the course—more than you thought you would, less than you thought or about as much as you thought?

6. What advice would you give an instructor who is using the networked classroom for the first time?

7. Did you use any programs or materials that were developed especially for the networked classroom? If so, how long did they take to develop; how much did they cost to develop?

8. Were any off-the-shelf computer programs used in this class? If so, which ones? Were they modified in any way for use in the class?

9. Would you like to teach another class in the NCS, or would you rather go back to a traditional classroom?

10. Did the students in your class display more creativity or better problem-solving skills and abilities than you had anticipated? If so, do you attribute this to use of the networked classroom?

11. Did you have any problems with the equipment or software in the networked classroom?[157]

A final methodology, the cognitive assessment of the students' performances, also examines the effectiveness of the NCS. Each instructor assesses the performance of his or her students during and at the end of the course. The assessments can take the form of exams, class projects, papers, or any other tool the instructor desires, and the instructor uses them to determine the extent to which the students have met the course objectives. Identical courses taught in traditional classrooms serve as control groups so evaluators can compare the students' performances in both learning environments. If no control group is available, then the evaluators can assess the results qualitatively to determine if they are significant.[158]

Even though the NCS assessment plan uses varied and extensive methodologies, problems do exist. For example, the student and instructor questionnaires are biased towards the networked classroom. In addition, the respondents to the questionnaires compare a specific application of the NCS to general uses of traditional classrooms; however, both students and instructors have had diverse experiences—both good and bad—in traditional classrooms. The issue, then, is to which traditional classroom are they comparing the NCS?

The end-of-course surveys the students complete also are suspect. The standardized forms possess the disadvantages previously described. In addition, while evaluators possess a large data base of student surveys that the Air Force Academy has compiled from earlier semesters, they cannot compare the data directly to the information collected from the students who take courses in the NCS. If instructors properly use the technology in the NCS, the courses that instructors have taught in the past differ from those they teach in the NCS. Further, to compare courses taught in the traditional classroom realistically to those taught in the NCS, evaluators can use only those student surveys that correspond to identical courses taught by the same instructor.

Finally, cognitive assessments face the same problems. Evaluators can compare student performances in different learning environments only if they begin with identical course goals and objectives for those environments. Otherwise, they will have no common baseline to compare the students' performances.

Evaluators can make better use of the NCS assessment plan to determine ways to optimize the networked classroom. The information they gather with the different methodologies addresses such questions as: What is the best way to use the NCS?; How can the Air Force Academy make the NCS more effective?; and, What aspects of the NCS can the Air Force Academy improve? Evaluators can use the assessments to determine if the availability of the

NCS alters the pedagogy and causes instructors to operate differently. For example, research may show if the NCS alters the educational process significantly or how student and instructor perceptions and motivation use the NCS change over time.[159] Finally, researchers can use the NCS assessment plan to support instructors in developing evolutionary or revolutionary styles of teaching.

Summary

A distinct advantage, however, of the NCS assessment plan is that it uses a number of different assessment methodologies to build a data base to use in comparing styles of teaching. Any assessment of teaching styles must include diverse methodologies so educators can formulate a complete picture of the effectiveness of those teaching styles. The next chapter identifies guidelines educators can follow when they develop methodologies to thoroughly assess different styles of teaching.

Notes

1. Andrew V. Beale, "Are Your Students Learning What You Think You're Teaching?," *Adult Learning* 4, no. 3 (January/February 1993): 18–19.

2. K. Patricia Cross and Thomas A. Angelo, *Classroom Assessment Techniques*, Technical Report No. 88-A-004.0 (Ann Arbor, Mich.: Board of Regents of the University of Michigan for the National Center for Research to Improve Postsecondary Teaching and Learning, 1988), 2.

3. Beale, 19.

4. Douglas R. Whitney, "Performance Assessment: Lessons From History," *Adult Learning* 4, no. 3 (January/February 1993): 23.

5. Marc J. Rosenberg, "Evaluating Training Programs for Decision Making," in *Evaluating Business Industry Training*, ed. Leslie S. May, Carol A. Moore, Stephen J. Zammet (Boston: Kluwer Academic Publishers, 1987), 58.

6. Ibid., 63.

7. Grant Wiggins, "Creating Tests Worth Taking," *Educational Leadership* 49, no. 8 (May 1992): 26–32.

8. Joseph P. McDonald, "Three Pictures of an Exhibition: Warm, Cool, and Hard," *Phi Delta Kappan* 74, no. 6 (February 1993): 484–85.

9. Irwin L. Goldstein and Associates, *Training and Development in Organizations* (San Francisco, Calif.: Jossey-Bass Publishers, 1989), 136–86.

10. Irwin L. Goldstein, "Content Validity as an Evaluation Strategy for Examining Training Programs," in *Evaluating Business and Industry Training*, ed. Leslie S. May, Carol A. Moore, and Stephen J. Zammet (Boston: Kluwer Academic Publishers, 1987), 139–41.

11. Susan Loucks-Horsley, "Science Assessment: What Is and What Might Be," *Educational Leadership* 46, no. 7 (April 1989): 86–87.

12. Goldstein, "Content Validity as an Evaluation Strategy for Examining Training Programs," 142.

13. Ibid., 146–47.

14. Rosenberg, 60–62.

15. Karen S. Brethower and Gary A. Rummler, "Evaluating Training," *Training and Development Journal* 33, no. 5 (May 1979): 14–22; Mary E. Diez and C. Jean Moon, "What Do

We Want Students to Know? . . . and Other Important Questions," *Educational Leadership* 49, no. 8 (May 1992): 38–41.

16. Michael Q. Patton, *Creative Evaluation* (Beverly Hills, Calif.: Sage Publications, 1981), 84–85.

17. Ibid., 29–49.

18. Beale, 19.

19. Cross and Angelo, 4.

20. Richard J. Stiggins, "Assessment Literacy," *Phi Delta Kappan* 72, no. 7 (March 1991): 536.

21. Ibid.

22. W. Edward Moomaw, "Practices and Problems in Evaluating Instruction," *New Directions for Higher Education* 5, no. 1 (Spring 1977): 85–88.

23. Arthur M. Cohen and Florence B. Brawer, *Measuring Faculty Performance* (Washington, D.C.: American Association of Junior Colleges, 1969), 62–65.

24. Robert L. Linn, "Measuring Pretest-Posttest Performance Changes," in *Educational Evaluation Methodology: The State of the Art*, ed. Ronald A. Berk (Baltimore, Md.: Johns Hopkins University Press, 1981), 85–105.

25. Ibid., 88.

26. Ibid.

27. Ibid., 89–95.

28. Richard M. Wolf, "Selecting Appropriate Statistical Methods," in Berk, 110–23; Goldstein and Associates, *Training and Development in Organizations*, 90–109.

29. Ibid.

30. Ibid.

31. Michael Q. Patton, *Utilization-Focused Evaluation* (Beverly Hills, Calif.: Sage Publications, 1978), 152–56.

32. David W. Dalton, Michael J. Hannafin, and Simon Hooper, "Effects of Individual and Cooperative Computer-Assisted Instruction on Student Performance and Attitudes," *Educational Technology Research and Development* 37, no. 2 (1989): 18–19.

33. Dr Michael Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), interview with author, 19 January 1993.

34. David W. Dalton and Michael J. Hannafin, "The Effects of Computer-Assisted and Traditional Mastery Methods on Computation Accuracy and Attitudes," *The Journal of Educational Research* 82, no. 1 (September/October 1988): 27–32.

35. Linn, 91–98.

36. Grant Wiggins, "Teaching to the (Authentic) Test," *Educational Leadership* 46, no. 7 (April 1989): 41–47; Lorrie A. Shepard, "Why We Need Better Assessments," *Educational Leadership* 46, no. 7 (April 1989): 4–9.

37. Blaine R. Worthen, "Critical Issues That Will Determine the Future of Alternative Assessment," *Phi Delta Kappan* 74, no. 6 (February 1993): 446.

38. Randy Elmore, "Improving the Assessment of Georgia Students," *Education* 113, no. 1 (Fall 1992): 108.

39. E. Peter Volpe, "The Shame of Science Education," *American Zoologist* 24 (1984): 437.

40. Shepard, 4–9.

41. Beale, 25.

42. Ibid.

43. Grant Wiggins, "A True Test: Toward More Authentic and Equitable Assessments," *Phi Delta Kappan* 70, no. 9 (May 1989): 708.

44. Noel Entwistle, "Teaching and the Quality of Learning in Higher Education," in *Handbook of Educational Ideas and Practices*, ed. Noel Entwistle (London: Routledge, 1989), 679; Elmore, 109; Worthen, 450.

45. Jon J. Denton, "Selecting an Appropriate Grading System," *The Clearing House* 63, no. 3 (November 1989): 107.

46. Sally Brown, "Criterion-Referenced Assessment," in *Handbook of Educational Ideas and Practices*, ed. Noel Entwistle (London: Routledge, 1990), 867–68.

47. Alistair Pollitt, "Diagnostic Assessment Through Item Banking," in *Handbook of Educational Ideas and Practices*, 875–77.

48. Ibid., 877–85.

49. Denton, 108–10.

50. Wiggins, "A True Test," 705–6; Maria Valeri-Gold, James R. Olson, and Mary P. Deming, "Portfolios: Collaborative Authentic Assessment Opportunities for College Developmental Learners," *Journal of Reading* 35, no. 4 (December 1991/January 1992): 298–99.

51. Brown, 866.

52. Ibid., 867–69.

53. Beale, 25.

54. Ibid.

55. George F. Madaus and Thomas Kellaghan, "The British Experience with 'Authentic' Testing," *Phi Delta Kappan* 74 no. 6 (February 1993): 459–69.

56. Ibid., 459.

57. Ibid., 459–69, Brown, 871.

58. Madaus and Kellaghan, 459–69.

59. Brown, 871.

60. Tom Bryce and Jim McCall, "Assessment of Practical Skills," in *Handbook of Educational Ideas and Practices*, 887–95.

61. Jan Heckley Kon and Giselle O. Martin-Kniep, "Students' Geographic Knowledge and Skills in Different Kinds of Tests: Multiple-Choice Versus Performance Assessment," *Social Education* 56, no. 2 (February 1992): 96.

62. Wiggins, "A True Test," 706.

63. Ibid., 711–12; Elmore, 109.

64. Elmore, 110.

65. Beale, 26.

66. Ibid.

67. Shepard, 4–9.

68. Sheila W. Valencia et al., "Theory and Practice in Statewide Reading Assessment: Closing the Gap," *Educational Leadership* 46, no. 7 (April 1989): 58; Allan A. De Fina, Linda L. Anstendig, and Kathryn De Lawter, "Alternative Integrated Reading/Writing Assessment and Curriculum Design," *Journal of Reading* 34, no. 5 (February 1991): 354.

69. Valencia et al., 59.

70. Ibid., 59–60.

71. Roger Farr and Beverly Farr, "Integrated Assessment System," *Journal of Reading* 35, no. 5 (February 1992): 417.

72. Valencia et al., 59–60.

73. Ibid., 60–61.

74. Ibid., 61–62.

75. Lt Col William Newmiller, Department of English, Air Force Academy (HQ USAFA/DFENG), interview with author, 11 August 1992.

76. Winfield Cooper and B. J. Brown, "Using Portfolios to Empower Student Writers," *English Journal* 81, no. 2 (February 1992): 40–45; Valeri-Gold, Olson, and Deming, 299–304.

77. Cooper and Brown, 40–45.

78. Newmiller interview.

79. Carolyn Hutchinson, "Communicative Competence in Language Teaching," in *Handbook of Educational Ideas and Practices*, 755.

80. Joseph Abruscato, "Early Results and Tentative Implications from the Vermont Portfolio Project," *Phi Delta Kappan* 74, no. 6 (February 1993): 475–76.

81. Whitney, 23.

82. John O'Neil, "Putting Performance Assessment to the Test," *Educational Leadership* 49, no. 8 (May 1992): 15.

83. Beale, 26.

84. Heckley Kon and Martin-Kniep, 98.

85. National Council for the Social Studies, "Testing and Evaluation of Social Studies Students," *Social Education* 55, no. 5 (September 1991): 284–86.

86. Ibid.; Michael J. Feuer and Kathleen Fulton, "The Many Faces of Performance Assessment," *Phi Delta Kappan* 74, no. 6 (February 1993): 478.

87. Richard J. Shavelson and Gail P. Baxter, "What We've Learned About Assessing Hands-on Science," *Educational Leadership* 49, no. 8 (May 1992): 20.

88. Ibid.

89. Ibid., 22–25; Loucks-Horsley, 86–87.

90. Walter Szetela and Cynthia Nicol, "Evaluating Problem Solving in Mathematics," *Educational Leadership* 49, no. 8 (May 1992): 42–45; Volpe, 437–39.

91. Volpe, 437–39.

92. Whitney, 24; Worthen, 452.

93. O'Neil, 18–19.

94. Anthony S. Bryk and Stephen W. Raudenbush, "Application of Hierarchial Linear Models to Assessing Change," *Psychological Bulletin* 101, no. 1 (January 1987): 147–48.

95. Ibid., 147–56.

96. Cross and Angelo, 22–23.

97. Ibid., 34–35.

98. Ibid., 30–32.

99. Ibid., 39–40.

100. Ibid., 52–53.

101. Ibid., 78–81.

102. Ibid., 66.

103. Ibid., 66–68.

104. Loucks-Horsley, 86–87.

105. Robert D. Abbott et al., "Satisfaction With Processes of Collecting Student Opinions About Instruction: The Student Perspective," *Journal of Educational Psychology* 82, no. 2 (June 1990): 201.

106. Robert T. Backburn and Mary Jo Clark, "An Assessment of Faculty Performance: Some Correlates Between Administrator, Colleague, Student and Self-Ratings," *Sociology of Education* 45, no. 2 (Spring 1975): 249; Cohen and Brawer, 10.

107. John A. Centra, "The How and Why of Evaluating Teaching," *New Directions for Higher Education* 5, no. 1 (Spring 1977): 93–106; Backburn and Clark, 249–50.

108. Herbert W. Marsh and Michael Bailey, "Multidimensional Students' Evaluations of Teaching Effectiveness: A Profile Analysis," *The Journal of Higher Education* 64, no. 1 (January/February 1993): 10–11.

109. Ibid., 1.

110. Cohen and Brawer, 11–12; Backburn and Clark, 250; Centra, 97–99.

111. Cohen and Brawer, 6.

112. Ibid., 7.

113. Ibid., 6–7.

114. Stephen E. Newstead and John Arnold, "The Effect of Response Format on Ratings of Teaching," *Educational and Psychological Measurement* 49, no. 1 (Spring 1989): 33–43.

115. Ibid.

116. Abbott et al., 202.

117. Naomi Karten, "Evaluations That Rate a '10': Are You Getting the most from Your Best Source of Student Feedback?" *Computerworld* 24, no. 7 (12 February 1990): 116.

118. David L. Silvernail and Judith L. Johnson, "The Impact of Interactive Televised Instruction on Student Evaluations of Their Instructors," *Educational Technology* 32, no. 6 (June 1992): 48–49.

119. Cheri Ostroff, "Training Effectiveness Measures and Scoring Schemes: A Comparison," *Personnel Psychology* 44, no. 2 (Summer 1991): 361.

120. Patricia Cranton and Ronald A. Smith, "Reconsidering the Unit of Analysis: A Model of Student Ratings of Instruction," *Journal of Educational Psychology* 82, no. 2 (June 1990): 208–9.

121. Herbert W. Marsh, "Multidimensional Students' Evaluations of Teaching Effectiveness: A Test of Alternative Higher-Order Structures," *Journal of Educational Psychology* 83, no. 2 (June 1991): 285–96.

122. Ibid.

123. Marsh and Bailey, 11–13.

124. Ostroff, 365–69.

125. Marsh, 285–96.

126. Cranton and Smith, 207.

127. Ibid., 208–11.

128. Gregory A. Jackson, "Evaluating Learning Technology: Methods, Strategies, and Examples in Higher Education," *The Journal of Higher Education* 61, no. 3 (May/June 1990): 298–99.

129. Linda Costigan Lederman, "Assessing Educational Effectiveness: The Focus Group Interview as a Technique for Data Collection," *Communication Education* 38, no. 2 (April 1990): 118.

130. Ibid., 117–19.

131. Ibid., 120.

132. Ibid.

133. Ibid., 120–26.

134. Abbott et al., 201.

135. Ibid., 201–2.

136. Ibid., 202.

137. Ibid., 204–6.

138. Jackson, 298–301.

139. Cross and Angelo, 138–39.

140. Ibid., 129–30.

141. Ibid., 148–49.

142. Ibid., 134–35.

143. Cynthia A. Lombardo, "Do The Benefits of Training Justify the Costs?", *Training and Development Journal* 43, no. 12 (December 1989): 60–64.

144. Nancy D. Padak and Gary M. Padak, "What Works: Adult Literacy Program Evaluation," *Journal of Reading* 34, no. 5 (February 1991): 374.

145. Goldstein and Associates, 65–86.

146. Ibid., 85.

147. Anthony P. Carnevale and Eric R. Schultz, "The Consensus Accounting Model," *Training and Development Journal* 44, no. 7 (July 1990): S-9 through S-10.

148. Ibid., S-10 through S-11.

149. Evelyn H. Nunes, "Evaluating Adult Instructional Materials," *Adult Learning* 3, no. 2 (October 1991): 16.

150. Ibid.

151. Ibid., 16–17.

152. Dan Surry, graduate assistant, Directorate of Education, Air Force Academy (HQ USAFA/DFER), interview with author, 11 August 1992; notes prepared by Dr Mary Marlino and Dan Surry for the Directorate of Education, July–August 1992.

153. Marlino and Surry notes.

154. Ibid.

155. Ibid.

156. Ibid.

157. Ibid.

158. Ibid.

159. Dr Michael Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, Air Force Academy (HQ USAFA/DFBL), interview with author, 12 August 1992.

Chapter 5

Recommended Guidelines for Assessment Methodologies

The previous chapter asserted that educators must use diverse assessment methodologies to assess the effectiveness of a teaching style. Indeed, according to Mary Marlino,

> Human learning is too complex to be adequately or meaningfully explained by isolated experimentally controlled variables. Few evaluation designs take into account a sufficient number of the salient variables involved in the entire learning process or the environment in which learning occurs.[1]

While no single methodology completely assesses the effectiveness of a teaching style, educators must identify specific parameters they require for a successful assessment. Evaluators use these parameters to ensure that their assessments of teaching styles are thorough and accurate, since the assessments must provide "decision-makers with timely, accurate information upon which to base educational decisions; it is not enough to know that a [style of teaching] is effective; decision-makers need to know why it is effective."[2]

Unfortunately, few assessment methodologies account for a large number of the salient variables associated with the complex nature of learning or styles of teaching, and evaluators usually manipulate only one or two of the variables to generate conclusions about the effectiveness of the teaching styles. Further, assessments typically treat teaching as a uniformly administered experience for students; in reality, learning is a different experience for each student.

A models of learning approach is one way to address these two issues. Models can account for a wide variety of educational variables, including student aptitude and motivation, student progress and achievement, and characteristics of the teaching style and the learning environment. Models can provide decision makers, instructors, and evaluators with accurate information about the effectiveness of different teaching styles.[3]

Models of Learning

A general model of learning in a classroom environment, the Carroll model assumes that five basic categories of variables account for differences in student achievement: aptitude, opportunity to learn, perseverance, quality of

instruction, and ability to understand instruction. The emphasis in this model is on time, as the first three variables are expressed in terms of time. Aptitude measures the time a student needs to learn a task; opportunity to learn indicates the amount of time the educational institution allows for learning; and perseverance is the amount of time a student is willing to engage in active learning.[4]

The other two variables in the model relate to student achievement. Quality instruction requires that teachers state clearly what students are to learn and ensures that students have access to the appropriate educational materials and that the learning process is carefully planned, organized, and executed. Decreasing the quality of instruction increases the amount of time students need for learning. Similarly, the time a student needs for learning depends on his or her ability to understand the instruction. This ability includes the student's language comprehension and capability to define, understand, and accomplish a particular learning task. Considering that time is an important variable in the Carroll model, the degree of learning, or achievement, is assumed to be a ratio of the time actively spent on learning to the time required to learn.[5]

Evaluators use another model—the causal model—to make logical inferences based on their gathered assessment data. They first establish a number of presumed, causal relations among the salient variables in the style of teaching and the learning environment. The evaluators then assess the model against the gathered data by using a correlation matrix that expresses observed relations among the variables corresponding to the data collected. A causal model normally assumes that two or more independent variables influence the same dependent variable; it can provide evidence of intervening variables that interact with, and are hence presumed to explain, the relationships of the independent and dependent variables.[6]

Evaluators have conceptualized models to account for the complexities of such innovative educational technologies as computer-assisted teaching and interactive videodiscs. One model consists of six levels designed to explain the information gathered from evaluations of these educational technologies. The levels are project documentation, assessment of the value of the course objectives, formative assessment, immediate effectiveness assessment, impact assessment, and cost-effectiveness assessment.[7]

The modeling techniques described in the preceding paragraphs address theoretical research issues more directly than traditional data analysis techniques. These models force the evaluator to conceptualize systems of variables rather than just isolated pairs of variables. Evaluators confirm a model by determining if it adequately describes the collected data; confirmation does not imply proof, nor does it validate the model.[8]

Still, the modeling approach to assessment potentially allows educators to examine a program in terms of student characteristics and implementation variables and can therefore assist them to determine the effectiveness of various teaching styles.[9] However, developing and validating models are time-intensive processes, and they require dedicated efforts from experts in

the field. Further, each model is individualized; that is, one model cannot adequately describe courses from diverse educational disciplines. For example, partial least squares, a form of causal modeling, was "developed specifically for the social and behavioral sciences."[10] Finally, evaluators cannot prove a model's validity. As Mary Marlino states, "Evidence simply cannot validate a theory. One can often find another theory that explains a given set of data, and the task of the researcher is to challenge and ultimately overturn our useful but limited theoretical models."[11]

Assessments Based on Student Achievement

Evaluators can design adequate assessment methodologies to be less involved than constructing various models of diverse teaching styles or learning environments and then fitting data to conform to the models. These alternative methodologies should be diverse to account for the complexities of learning, but they also should be rather simple to implement. Student achievement of learning objectives should be the main criterion upon which educators base their studies of teaching effectiveness. They should examine each student's performance individually by determining the effectiveness of teaching specific objectives to specific types of students.[12]

For example, low-ability students may have high gain scores; they could have very low pretest scores and average posttest scores. Conversely, high-ability students may have high scores on both their pre-test and posttest and, therefore, low gain scores. While both groups of students may achieve the course objectives, some educators argue that the low-ability students outperformed the high-ability students based on the results of the gain scores. On the other hand, the high-ability students may have higher posttest scores and, hence, possess more thorough knowledge and understanding of the course objectives. These students, therefore, outperformed the lower ability students in the eyes of other educators. Consequently, evaluators assessing the effectiveness of the teaching style cannot examine only gain scores or posttests. Rather, they must consider the performance of the low- and high-ability students separately before making definitive conclusions concerning a teaching style's effectiveness.

If student achievement is indeed the primary criterion on which to assess the effectiveness of teaching styles, then an assessment methodology should possess the following attributes:

1. Educators administer pretests, questionnaires, and surveys to all students involved in the assessment. These tools establish a baseline for assessing learning.

2. Instructors teach the students identical subject matter to insure a commonality of learning objectives.

3. Students periodically complete tests, questionnaires, and surveys during the course to help evaluators track their evolving knowledge, skills, abilities, understanding, and comprehension.

4. Educators administer posttests to all students. Multiple tests assess the students' behavioral and cognitive achievements.

5. At the end of the course the students complete questionnaires and surveys to determine how their attitudes, opinions, and motivation levels have changed since the course began.

6. Students rate the effectiveness of the teaching style.

7. Fellow instructors periodically judge the effectiveness of the teaching styles based on factors other than what the students have learned.[13]

As the above attributes suggest, while student achievement plays a central role in determining the effectiveness of different teaching styles, it must not be the only criterion. Indeed, the mere acceptance of test scores may not accurately portray the learning that occurs within a classroom environment; the results may be misleading or even useless. Gary Negin concludes that,

1. testing is sometimes conducted before clear, worthwhile goals are identified

2. higher levels in a bureaucracy do not always communicate their intentions and purposes to those below

3. testing conditions are sometimes violated

4. tests and results are sometimes misused

5. potential injuries to students [and instructors] can occur

6. numbers do not tell the whole story.[14]

According to Andrew Beale, "A variety of student assessment data yields a more vivid and reliable picture of student growth and learning, plus such information serves as the basis for decisions about curriculum and [studies of teaching] and overall program integrity."[15]

Common Characteristics of Assessments

Educators must therefore use multiple criteria to assess the effectiveness of a teaching style accurately. Richard Clark offers several recommendations to the educators who identify these assessment criteria:

1. when comparing different [styles of teaching], insure that controls for the [teaching style], student contact time, and curriculum content are in place

2. focus on measuring "cost," like time savings, student access, dollar investment, and maintenance, rather than on achievement gains

3. discourage [assessments] of programs using student volunteers; instead assignment to the programs should be based on ability and/or prior knowledge of the subject matter

4. consider administering a pretest, as well as a survey of student preferences and beliefs about such program features as medium and [style of teaching]

5. be alert for achievement and preference antagonism due to interactions between student ability/prior knowledge differences and the [style of teaching]

6. generally, any medium can be used to deliver instruction at any level.[16]

Based on the above recommendations, any assessment methodology must possess several common characteristics. One key characteristic is the relationship of the methodology to the course and style of teaching it is designed to assess. Specifically, educators should develop the assessment methodology and the course simultaneously. Educators must first establish specific goals and objectives for the course so that they can determine the specific objectives of the assessment.[17] The assessments then reflect the stated objectives and the opportunities for student learning; what the curriculum instructors teach is the curriculum the evaluators assess. This curriculum encourages evaluators to use a variety of assessment tools to assess not only basic knowledge and skills but also higher level skills, abilities, and understanding.[18]

The educators must therefore form a small cadre who will develop the goals and objectives for both the course and the assessment. This cadre must include the instructors involved in the course, the evaluators, the decision makers, and the information users. The first two groups of people are easily identified, but identifying the decision makers and the information users may be more difficult. These two groups are comprised of, according to Michael Patton,

1. people who can use the information;
2. people to whom the information can make a difference;
3. people who have questions they need to have answered; and
4. people who care about and are willing to share the responsibility for the [assessment] and its utilization.[19]

Behavioral and Cognitive Objectives

Chapter 1 emphasized that the cadre must delineate specific, observable, and measurable behavioral and cognitive objectives which identify concrete outcomes of the course and definitive levels of student competency. The cadre must specify the desired objectives in terms of only those student behaviors and cognitions which the cadre can strictly define, observe, and measure quantitatively or qualitatively. The cadre also must define the desired levels of learning and the associated type of the learning—whether it is short-term performance in class or long-term transfer of knowledge, skills, abilities, and understanding to related courses. The identified objectives and learning levels are then the basis for establishing the assessment objectives. Once the members of the cadre constrain the assessment by identifying the required objectives, they can develop the appropriate assessment methodology or methodologies.

Identifying the Students and Instructors

The cadre also should have the responsibility for identifying the students and instructors who will participate in an experiment to compare different styles of teaching. Since it is difficult to conduct a double-blind experiment where both the instructors and the students are unaware of their involvement in the experiment, a single-blind experiment is preferable; the instructors must know of the experiment, but the students don't have to know that they are participating in the experiment. An interesting study would compare students enrolled in a variety of courses that use diverse traditional teaching methods and different innovative educational technologies. For now, assume that the cadre will compare two styles of teaching applied to a single course. In one class the instructor will use a single traditional teaching method and in the other class the instructor will teach the students by using one innovative educational technology as a teaching aid. After establishing a single set of course goals and objectives and developing two distinct learning environments, the cadre needs to equalize, between groups, the characteristics of the students chosen to participate in the experiment.

Note that the cadre must choose students for the study. The students should not volunteer because they would introduce unavoidable effects which may bias the results of the experiment.[20] Students differ from one another in the knowledge, skills, abilities, understanding, and attitudes they bring into the learning environment. To a great extent these individual characteristics determine students' motivational levels and performances in class. Hence, the cadre needs to identify the students and, where possible, place them in the two classes before the course begins. If the members of the cadre cannot identify students until the beginning of the course, they should initially bring the students together in a neutral environment, such as a lecture hall. After the members determine which classes to assign the students, they send the students to their respective classes without telling them that they are participating in an experiment.

Identifying the students is a rather difficult task, as it requires in-depth knowledge of each student. The cadre must establish a set of characteristics which adequately describes the students. Ideally, the cadre should possess an extensive background summary of each student, to include such general ability measures as intelligence quotient tests and grades in prior courses. The cadre should give students one or more pretests to establish their prior knowledge and understanding of the subject. In addition, the evaluators should use the pretests to establish a baseline to compare the scores from tests the students take during the course. The cadre also should develop and administer one or more thorough questionnaires and surveys to discriminate such student characteristics as

1. age, sex, and race;
2. interest in the particular subject matter;
3. motivation towards learning the subject matter;
4. approach and attitude towards learning;

5. desire to learn;
6. attitudes towards traditional teaching methods and innovative educational technologies;
7. preferred style of learning;
8. special abilities and competencies;
9. aptitude towards computer usage;
10. course load;
11. available time for learning;
12. quality of study time; and
13. outside sources of stress.

The cadre members must quantify these variables so that they can divide the students into two identical groups. Although the surveys and questionnaires will be extensive, the cadre cannot identify and quantify all student characteristics. Thus, the cadre members must clearly delineate those characteristics that they cannot identify and quantify. To determine if these variables differentially and systematically influence a student's performance, the cadre members must accomplish three tasks. First, they need to impanel experts to judge the results of the experiment to determine their freedom from possible bias introduced by unidentified characteristics. Second, evaluators who analyze the data from the study should employ statistical techniques to measure the resulting bias and then develop correcting factors. Third, the members of the cadre must limit their interpretations of the results to only those justified by strong and valid evidence.[21]

Based on the data from the pretests, surveys, and questionnaires, the cadre can form a control group and an experimental group. The students in the control group take the class an instructor teaches in the traditional manner and the students in the experimental group enroll in the class where an instructor uses the innovative educational technology as a teaching aid. The cadre averages the quantified characteristics of each student within each group to form a composite student; the composites must be nearly identical between groups.

Ideally, the cadre should characterize the instructors for the two classes in a similar manner. Because an instructor can influence a student's performance in class, the cadre must quantify an instructor's characteristics where possible. For example, instructors possess a wide range of teaching experience and confidence levels, and the cadre must quantify these two characteristics. In addition, the cadre must quantify the instructors' attitudes and opinions towards teaching. Surveys, questionnaires, and classroom observations are effective tools to identify each instructor's characteristics before the experiment is conducted; chapter 4 discusses the questions which the cadre can use to develop the appropriate questionnaires and surveys.

However, the members of the cadre must consider a number of variables as they identify the characteristics of the instructors. For example, the students' attitudes towards the instructor, the course, and the learning process influence an instructor's perceptions of the students and his or her attitudes

towards the students and the course. The instructor's experience, confidence, and attitude necessarily influence student performance. In addition, the instructor may exhibit bias towards a particular teaching style, and this bias may influence the students' performances in class. Such variables affect the instructor's attitudes and performance in the classroom and, therefore, his or her characteristics with respect to the teaching skills and abilities he or she possesses.

While the members of the cadre will find it difficult to quantify these variables, they must specify as many characteristics as possible to identify the instructors for the experiment. They should request volunteers to teach the classes by using the two different styles of teaching. Nonvolunteers who have negative attitudes towards participating in the study may become biased as they teach. The cadre members should choose the volunteers from a pool of instructors they have identified based on their specified characteristics.

One dilemma the cadre faces is to identify the types of instructors required to teach the course in the traditional manner and in an innovative educational technology setting. The cadre members must first determine whether a single instructor should teach the course in both learning environments or whether one instructor should teach one group of students with the traditional teaching method, while another instructor teaches the other group of students using the innovative educational technology as a teaching aid. Both options have the potential to introduce bias into the experiment. The single instructor may be better qualified to teach in one environment, and this qualification may influence his or her teaching of the students in the other environment. Similarly, if the cadre uses two instructors, their experience levels and attitudes may differ, thereby influencing their teaching skills and abilities. Students in one group may be at a disadvantage if they have an instructor with less experience and confidence than the other instructor.

Dan Surry has proposed an interesting alternative to the above two options. He designed a technological intervention assessment specifically to determine the impact of a video-based presentation system for the Political Science Department at the US Air Force Academy. The assessment dictates that the instructor teaches both video-based and traditional lectures. Evaluators select four instructors for the experiment. These instructors possess varying levels of teaching experience and experience using innovative educational technologies. One instructor should have experience with traditional teaching methods but little experience with innovative educational technologies. Another instructor should have little experience in the traditional classroom but should be experienced in using innovative educational technologies. A third instructor should have some experience with teaching in the traditional classroom and in using innovative educational technologies. The fourth instructor should be quite experienced both in teaching in the traditional setting and in using innovative educational technologies.[22]

If there are enough instructors available, and if there are enough students enrolled to form eight classes for the experiment and at least one control

class, then this proposal becomes intriguing. Researchers then can conduct a study to determine how instructors' attitudes towards different styles of teaching and experience levels in teaching influence student performance. Evaluators may uncover possible biases in the instructors' attitudes if they track instructors with different experience levels who teach the same course in both learning environments.

Unfortunately, the cadre members may not have the luxury of a pool of instructors with varying degrees of experience with traditional teaching methods and in teaching using innovative educational technologies. A compromise experimental setup could utilize two instructors who are experienced teachers of the subject matter the students will be exposed to in the experiment. The first instructor would teach the control group and use the traditional method of teaching the cadre chose for the experiment. This instructor should be an experienced implementer of this teaching style and have a positive attitude towards the style and the learning process itself. The second instructor would teach the experimental group using an innovative educational technology as a teaching aid. He or she must be an experienced user of innovative educational technologies and have a positive attitude towards using the technologies in the classroom environment. The positive attitudes of both instructors should propel them to teach their best, and any instructor bias which may influence student performance should be minimized. The two instructors would work together to develop the course, ensuring that they would meet all of the identified behavioral and cognitive objectives. However, each instructor would develop their respective lessons in ways that optimize their teaching styles and their learning environments.

Once the members of the cadre select the course, instructors, and students for the experiment, they can begin the experiment; in other words, the instructors can begin teaching the course. At this stage evaluators can begin to conduct the actual assessments. Recall that the students already have completed at least one pretest, and evaluators can use this pretest as a baseline to assess the students' progress towards attaining the course objectives. While every course has different assessment requirements, it should have a methodology that possesses numerous assessments conducted throughout the course. These assessments can take many forms, and the cadre must understand the purposes and advantages of each.

Types of Assessments

Norm-referenced assessments of learning are usually quantitative, or behavioral, measures of student performance in a learning environment. Such norm-referenced assessments as standardized or multiple-choice tests are necessary to test student knowledge of basic skills and factual recall. In addition, these assessments indicate a student's performance as compared to

others. Item banking, as a norm-referenced assessment, is a valuable tool for establishing a scale with which to compare student performance.

Behavioral Measures of Learning

These behavioral measures of learning usually rely on gain scores based on each student's pretest results; chapter 4 describes several methods of calculating various types of gain scores. Probably the most straightforward technique estimates true change in a student's test scores by using regression analysis to calculate the gain scores. If reliability estimates of the pretests and posttests are not available, the standardized change score is an alternative methodology to use. However, evaluators assume student characteristics that may influence student performance to be similar, which may be an invalid assumption in many situations. Evaluators can implement such complex analysis procedures as the analysis of covariance if they are familiar with those tools.

One difficulty with using gain scores as an assessment tool is that evaluators use pre-tests and posttests as the sole measure of the students' performances, and they fail to consider the students' performances in between the tests. One possible solution to this problem is to administer multiple tests throughout the course. Evaluators also can use multiple measures of student performance to track each student's behavioral growth during the course. However, when designing the course and the associated tests, the instructors and the evaluators must ensure that they do not design the course to teach the test.

Cognitive Measures of Learning

While many courses require student knowledge of specific learning outcomes, they also require the students to understand and master underlying concepts, skills, and abilities related to the course. Therefore, instructors and evaluators must design criterion-referenced, authentic, and performance assessments that require more than recall of material by testing students on their ability to analyze and synthesize information. These assessment methodologies challenge students to extend their knowledge and understanding of a topic. The cadre should not reference the cognitive assessments of one student's performance to others; rather, the assessments permit the cadre to determine how effectively each student achieves the lesson objectives. Chapter 4 briefly summarizes two examples of criterion-referenced assessments; it also discusses authentic and performance assessments which emphasize realistic complexity and stress depth more than breadth of knowledge. Although the process of developing, administering, and critiquing authentic or performance assessments is a difficult and complex proposition, the cadre should use them as part of their methodology when they assess the effectiveness of different teaching styles.

Unfortunately, many of the assessment methodologies the cadre can use to analyze student performance require quantifiable data. Important aspects of

student performance may not be quantifiable; hence, the cadre must consider alternative assessments to assess those qualitative aspects of student performance. Again, chapter 4 reviews several such assessments. The cadre members must ensure that the methodologies they employ match the style of teaching and assess student achievement of those desired learning objectives they have specified.

Student Evaluations of Teaching Styles

Though assessing students' performances in a learning environment is an essential part of any assessment methodology, it should not be the sole means of assessing the effectiveness of a teaching style. Student evaluations provide the cadre with important data for a comprehensive assessment. Implementing these evaluations requires that educators develop several additional components to the assessment methodology.

Student evaluations are effective tools to help the cadre assess the effectiveness of a teaching style. The members of the cadre already have developed questionnaires and surveys to acquire information on students so that they may place the students in the appropriate groups for the experiment. Cadre members can modify these forms to determine student opinions and attitudes during and after the experiment. They should administer the evaluation forms to the students to complete several times during the course. The cadre not only will gather essential information concerning the students' changing opinions and attitudes, but the members probably will discover that students take the forms more seriously if they perceive that the cadre will consider their opinions and suggestions during the experiment.

The cadre must devote considerable attention to designing the questionnaires and survey forms, as the appearance of these forms influences the students' response. While the cadre members can use standardized forms, they should not rely on them as the sole tool to solicit student responses. In developing the standardized forms, the cadre needs to develop ratings scales and wordings which do not influence or bias student responses. Developing the questions themselves requires extensive thought. Chapter 4 lists types of questions that the cadre possibly can use. The more numerous the questions, the more thoroughly the cadre members can determine student opinions and attitudes. They should categorize these questions to simplify their analysis, but the categories should be extensive and descriptive of important characteristics of the teaching style to help the cadre members delineate its effectiveness.

In addition to completing the standardized forms, cadre members should encourage the students to submit forms in which they can express their opinions and attitudes in their own words. Chapter 4 describes several techniques to accomplish this task. To foster openness from the students, the cadre should use the techniques periodically during the semester. The

one-minute paper is an example of one survey which is quick and easy for the students to accomplish and for the cadre members to analyze. The cadre members can use it frequently, and they can ask questions which concentrate on specific characteristics of the teaching style they wish to assess. The small group instructional diagnosis offers another technique to solicit student opinions, but this technique is more time-intensive; and some students may not wish to participate in an oral discussion of their opinions.

Faculty Evaluations of Teaching Styles

In addition to soliciting student opinions and attitudes concerning the effectiveness of a teaching style, the assessment methodology should incorporate the opinions and attitudes of the instructors and their faculty colleagues. The cadre should encourage the instructors' colleagues to attend the classes periodically and to assess the style of teaching. These individuals must be volunteers who are trained evaluators, and they must visit the classes frequently. They should assess both types of teaching styles the cadre is examining in the experiment, and they should complete evaluation forms after each visit. These forms should include standardized forms, questionnaires, and surveys which solicit written feedback from the faculty members.

Instructors also can provide valuable feedback on the effectiveness of the teaching style they are using, although they tend to have erroneous perceptions of how others view them. Hence, instructors should not conduct self-evaluations; rather, the cadre should encourage them to provide objective assessments of how they perceived their teaching styles worked for specific lessons. Instructors involved in the experiment should complete daily logs which specify such information as effective teaching styles, uses of the classroom and the innovative educational technology, problems they encountered, perceptions of student interest and involvement, classroom activities, and their opinions of the success of the lesson. Periodically, the cadre should compare these logs to the students' one-minute papers to determine correlations between the perceptions and opinions of both the instructors and the students.

As an interesting supplement to the daily log, the cadre should periodically gather the instructors involved in the experiment and have them discuss their perceptions and opinions of the experiment and the two styles of teaching. Such a group interaction may illuminate problems associated with the experiment or concerns the instructors may have.

Finally, at the end of the experiment, each instructor should complete a survey which solicits written feedback on the experiment and the styles of teaching. This survey should not query the instructors' opinions of their performances; rather, it should focus on the instructors' objective opinions of the effectiveness of the teaching styles, what lessons they learned, and what

changes the cadre should make to the styles of teaching, the learning environments, the lessons, and the course itself.

Some Caveats on Assessing the Effectiveness of Teaching Styles

The information compiled from the assessments of the students' performances and the surveys and questionnaires the students, the instructors, and the instructors' colleagues complete could provide an extensive data base with which the cadre could begin to analyze the effectiveness of the teaching styles chosen for the experiment. I must emphasize that cadre members must not conduct any such experiment only once. Rather, they must conduct the experiment for many semesters to acquire a large enough data base to minimize unwanted bias from variables that they have not accounted for.

Further, the cadre must include several different courses from various disciplines in the experiment. It is quite feasible that the results exhibited in one course could differ significantly from the results exhibited in another course from a different discipline. The cadre members must track and periodically test the students involved in the experiment to determine their long-term retention of knowledge, skills, abilities, and comprehension. They also should survey the students to determine how their opinions and attitudes towards learning and the different styles of teaching change in subsequent courses. Similarly, cadre members must survey the instructors involved in the experiment some time after their involvement in the experiment to determine their changed opinions and attitudes. The cadre should involve the instructors in the experiment more than once, as their experience levels and attitudes towards the different teaching styles will surely change as they use those styles of teaching. However, the cadre members should encourage new instructors to participate in the experiment to broaden their data base and to provide the instructors with the opportunity to broaden their teaching experiences.

Summary

In sum, this experiment requires a long-term commitment from the cadre responsible for initiating it. The decision makers must be especially aware of the fact that any results from the experiment will take a long time to compile and then prepare. They also must realize that the assessment methodology associated with this experiment should not involve any cost evaluations. Not only is it difficult, if not impossible, to attach a price to assessing the effectiveness of a teaching style, but decision makers should not attempt to associate a price with the experiment. While the experiment is a physical

process with attached costs, the cadre should consider it as a long-term research project requiring funds for its operation. The cadre should specify the level of funds they require to conduct the experiment, and this specification should be the extent of any cost assessments of the teaching styles or, for that matter, the learning environments.

This chapter has outlined the essential requirements of any assessment methodology that educators can use to assess the effectiveness of different teaching styles. The specific components of the methodology will depend on the types of courses and styles of teaching that the educators have selected to assess. With any assessment methodology, the cadre must examine all facets of the course and teaching styles carefully to ensure that the methodology or methodologies are appropriate and realistic and that the cadre will accumulate valuable information. There is no one best methodology; rather, the members of the cadre must be imaginative and resourceful to develop the best methodology for their purposes. Chapter 6 provides basic principles which educators can follow when they develop a comprehensive methodology to assess the effectiveness of different teaching styles.

Notes

1. Mary R. Marlino, "A Proposal for the Utilization of a Process Modeling Approach in the Evaluation of Interactive Videodiscs" (paper presented at the annual meeting of the American Educational Research Association, Boston, 1990), 3.
2. Ibid., 5–6.
3. Ibid., 16.
4. John B. Carroll, "The Carroll Model: A 25-Year Retrospective and Prospective View," *Educational Researcher* 18, no. 1 (January/February 1989): 26–31.
5. Ibid.
6. Marlino, 26.
7. Thomas C. Reeves and R. M. Lent, "Levels of Evaluation for Computer-Based Instruction," in *Instructional Software: Principles and Perspectives for Design and Use*, ed. D. F. Walker and R. D. Hess (Belmont, Calif.: Wadsworth, 1984).
8. Marlino, 19–32.
9. Ibid., 32–33.
10. Ibid., 32.
11. Ibid., 19.
12. Arthur M. Cohen and Florence B. Brawer, *Measuring Faculty Performance* (Washington, D.C.: American Association of Junior Colleges, 1969), 57–65.
13. Ibid., 67.
14. Gary A. Negin, "What Test Results Don't Reveal," *The Clearing House* 63, no. 3 (November 1989): 124.
15. Andrew V. Beale, "Are Your Students Learning What You Think You're Teaching?" *Adult Learning* 4, no. 3 (January/February 1993): 26.
16. Richard E. Clark, "The Role of Media in the Evaluation of Training," in *Evaluating Business and Industry Training*, ed. Leslie S. May, Carol A. Moore, and Stephen J. Zammet (Boston: Kluwer Academic Publishers, 1987), 158–64.
17. Gregory A. Jackson, "Evaluating Learning Technology," *The Journal of Higher Education* 61, no. 3 (May/June 1990): 309.
18. J. Merrell Hansen, "Outcome-Based Education: A Smarter Way to Assess Student Learning," *The Clearing House* 63, no. 4 (December 1989): 172–73.

19. Michael Q. Patton, *Utilization-Focused Evaluation* (Beverly Hills, Calif.: Sage Publications, 1978), 284.

20. Lt Col Frederick Gibson, Department of Behavioral Sciences and Leadership, US Air Force Academy (HQ USAFA/DFBL), interview with author, 13 August 1992.

21. Nancy S. Cole and Anthony J. Nitko, "Measuring Program Effects," in *Educational Evaluation Methodology: The State of the Art*, ed. Ronald A. Berk (Baltimore: Johns Hopkins University Press, 1981), 45–50.

22. Dan Surry, graduate assistant, Directorate of Education, US Air Force Academy (HQ USAFA/DFER), interview with author, 11 August 1992; and notes prepared by Dr Mary Marlino and Dan Surry for the Directorate of Education, US Air Force Academy, July–August 1992.

Chapter 6

Principles to Guide the Development of Assessments

Chapter 5 developed guidelines educators can follow to develop methodologies to assess the effectiveness of teaching styles. I based these guidelines on the information contained in the first four chapters of this report, because they provide useful information for educators who wish to assess different styles of teaching. These guidelines are useful also for educators who develop new courses, since they must establish methodologies to assess the performances of the students who will enroll in those courses. In addition, the educators must determine the effectiveness of the courses so that they can enhance and refine the courses. They can use the guidelines in chapter 5 to help them develop methodologies to assess student performance in a learning environment and to assess the effectiveness of the teaching styles within that environment.

While the guidelines described in chapter 5 provide a road map for developing assessment methodologies, they also highlight basic principles to which educators must adhere if they desire accurate and effective methodologies. These principles are basic tenets educators should consider when they develop assessment methodologies.

Basic Assessment Principles

Assessments are essential to any course; they are not isolated components which measure student performance at discrete points in time. As such, educators must integrate them into the course development process. This development process should possess the following characteristics:

1. Educators must assemble a cadre to develop the course; the instructors who will teach the course, the evaluators of the course, the decision makers, and the information users comprise the members of this cadre.
2. The cadre must establish clear, realistic, achievable, and observable goals for the course.
3. Based on these goals, the cadre then must develop specific, realistic, observable, and measurable behavioral and cognitive objectives which they can assess.

4. The cadre also must define the associated levels of learning, and specify them as either short-term performance in class or long-term transfer of knowledge to related courses.

5. The cadre must develop appropriate methodologies to assess student achievement of the behavioral and cognitive objectives at the desired learning levels.

If the cadre developing the course wants to compare it to an existing course, they must make certain their comparison study is strictly controlled. Normally, the study should compare the results of the performances of students in control and experimental groups. Students in the control group take the existing course while students in the experimental group take the course which the cadre is developing and assessing. The cadre must adhere to several criteria:

1. Both courses must have identical goals, behavioral and cognitive objectives, and associated levels of learning.

2. The cadre selects students to participate in the study; they are not volunteers.

3. The cadre should equalize, between groups, those characteristics of the students the cadre has identified as essential to the study.

4. The cadre also must identify student characteristics which they cannot quantify through pretests, questionnaires, and surveys.

5. The cadre must select instructors to participate in the study and then identify and quantify their characteristics.

If the members of the cadre wish to design assessment methodologies to determine if a student attains the desired learning objectives at the appropriate levels of learning, they must first establish a baseline performance measure. The cadre will then assess the students based on that measure, whether it is a pretest, questionnaire, or survey. The assessment methodologies must be systematic, comprehensive, and employ both quantitative and qualitative assessments.

Student surveys and questionnaires are effective tools that help the cadre assess the effectiveness of a teaching style. The students need to complete these evaluation forms several times during the course, in addition to completing them at the beginning and at the end of the course.

The instructors of the course the cadre is examining should objectively assess the effectiveness of specific lessons in the course curriculum. They should not conduct self-evaluations; rather, the instructors should provide opinions of their perceptions of the teaching style.

The instructors' colleagues should assess the effectiveness of the style of teaching by periodically attending classes and completing evaluation forms. They should summarize their observations of the teaching style and the learning environment. The cadre needs to train these individuals as evaluators so they can provide objective opinions of the teaching style and the environment rather than of the instructors themselves.

Assessing a style of teaching requires a long-term commitment from the members of the cadre. They must conduct a course several times to determine its short- and long-term impact on student performance. The cadre members must accept the fact that while they can receive immediate feedback on certain aspects of the course, they cannot compile comprehensive data on the effectiveness of the teaching style if they assess the course only once or twice. An accurate and objective evaluation of a teaching style requires a large data base that consists of multiple assessments of different courses conducted over an extended period of time.

The cadre must assemble a blend of assessments to provide a complete picture of student performance and an accurate measure of the effectiveness of the teaching style. Test scores alone may not accurately portray the learning that occurs within a classroom environment; multiple and diverse assessments better reflect the learning process.

While the assessment methodologies should be diverse enough to account for the complexities of the learning process, they should be simple to implement.

It is difficult, if not impossible, to attach a price to student performance in a learning environment. If the main criterion of the assessment methodology is how well students achieve the learning objectives, then evaluations based on the costs associated with the learning environment and the benefits of that environment are difficult to quantify. Educators may express costs and benefits in terms of time savings, student access, and dollar investments; but they cannot directly associate these costs and benefits with student performance.

Summary

The principles outlined above can help educators develop methodologies to assess different styles of teaching. Because educators must integrate assessments into the course as it is developed, these principles also guide educators who are either building new courses or enhancing existing ones. Educators must address these basic principles if they wish to develop an effective course with the appropriate means of assessing not only the performance of the students but also the effectiveness of the course and teaching style.

A realistic application of these principles is appropriate to help educators better understand how they can develop a course and the associated assessment methodologies. A course development program at the Air Force Academy provides an ideal opportunity to apply the principles and to implement the guidelines described in chapter 5. Chapter 7 examines this program and uses the guidelines and principles established in chapters 5 and 6 to help instructors at the Air Force Academy to develop this course.

Chapter 7

Applying the Guidelines and the Principles of Assessments to a Case Study: Engineering Mechanics 120

The ideal conclusion to this report applies the guidelines established in chapter 5 and the principles identified in chapter 6 to a realistic scenario. For this report, such a scenario would use the guidelines and principles to help educators develop assessment methodologies for an undergraduate course. Because they should integrate the assessments into the course development process, educators would use guidelines and principles as they develop the course itself.

A course development program in existence at the Air Force Academy provides a perfect setting for such an application. However, before I actually apply the guidelines and principles to this program, I must first describe it and detail some background information.

Engineering Mechanics 120

The Department of Engineering Mechanics (DFEM) and the Directorate of Education (DFE) recently commenced a joint educational initiative to enhance the engineering mechanics 120 course. Titled "Fundamentals of Engineering Mechanics," it is the first engineering course cadets enroll in at the Air Force Academy. While this required core course introduces cadets to the engineering disciplines, "cadets enrolled in engineering mechanics 120 have a wide range of background experiences, technical skills, and career goals."[1] The course addresses topics germane to nearly all engineering fields; specifically, these topics focus on fundamentals of statics and strengths of materials. As such, the Air Force Academy departs from the introductory engineering courses found in traditional engineering programs; they usually teach these two fundamental engineering topics as separate, sequential, full semester courses.[2] The course also requires students to apply critical thinking in a technical arena, master problem-solving methods, and learn laboratory operations in the engineering mechanics laboratory. To achieve these goals, the cadets receive a mix of lectures, demonstrations, laboratory exercises, tours, and in-class board work.[3]

As innovative as engineering mechanics 120 is, DFEM continually seeks to enhance the course. During the fall 1991 and spring 1992 semesters, the department implemented a series of "Total Quality Education" (TQE) initiatives in several sections of engineering mechanics 120 as a controlled experiment. "These TQE

initiatives included cooperative learning, group exercises, immediate feedback on testing, and immediate review and retesting of unsatisfactory skills."[4] While the students exposed to this learning environment responded well to the group exercises, course time surveys indicated that these students were no more motivated than their peers who were enrolled in the traditional engineering mechanics 120 course. In addition, even though students in the experimental sections avoided cramming before exams, they failed to retain information longer than the students in the controlled classes.[5]

While the DFEM instructors were disappointed with the results of this study, they were still determined to enhance engineering mechanics 120. Therefore, DFEM decided to examine this course in a fresh light by not being constrained by the present course content. In fact, several instructors in DFEM recorded their thoughts about engineering mechanics 120, identified strengths and weaknesses of the current course, and presented issues the department must address before it could revise or enhance engineering mechanics 120. The instructors' opinions remain anonymous, but I have summarized them below:

1. We should not lose sight of our goal of teaching a permanent understanding of the basic fundamentals of engineering mechanics to all future Air Force officers.

2. Engineering mechanics 120 tends to focus on content as opposed to the objectives and the process of the course.

3. Engineering mechanics 120 has a built-in conflict by trying to be all things to all people. It is designed to teach the fundamentals of statics and strengths of materials for engineering students while providing nonengineering students a grounding in those two topics.

4. The purpose of engineering mechanics 120 is twofold: to prepare nontechnical majors for the other six required engineering courses and to provide engineering majors the fundamentals for follow-on engineering courses.

5. A core program should provide students a firm foundation for further learning, development, and achievement through the teaching, study, and research of a body of fundamental knowledge.

6. A core program should support a majors program by enabling students to apply fundamental principles to solving problems and to discover new knowledge in their selected areas of study.

7. The first engineering course in an academic curriculum should challenge all students, both those on technical tracks and those on nontechnical tracks.

8. Engineering mechanics 120 presents an opportunity for students to gain an appreciation of engineering and its limitations, capabilities, and importance rather than as an obstacle in their learning process.

9. Engineering mechanics 120 should use and integrate math and science skills and knowledge and apply them to basic but meaningful engineering problems and projects.

10. Engineering mechanics 120 should provide a smooth transition from high school level to college level technical subjects.

11. Engineering mechanics 120 should have a strong mechanics thrust to connect science and math with engineering. It should cover basics of statics,

strengths of materials, kinematics, and kinetics, as well as an introduction to the characteristics of different materials. The course should allow students to establish a temporary working knowledge in those engineering topics.

12. Engineering mechanics 120 should provide a positive introduction to engineering and highly motivate cadets towards engineering careers.

13. Engineering mechanics 120 should provide students with hands-on design experiences so they develop a "feel" for engineering.

14. Engineering mechanics 120 should introduce a wide range, but a narrow depth, of engineering fundamentals.

15. Engineering mechanics 120 currently focuses on relatively few areas; it has virtually no design content and is highly mathematical in its approach.

16. Engineering mechanics 120 should introduce a plethora of simple engineering concepts; throughout the course students should design simple projects and have DFEM personnel build them. The students would test the finished products, and they would redesign failures.

17. Students need to appreciate the complexities involved in the engineer- ing design process and have the opportunity to apply engineering principles to a real-life design. This would help them truly understand those principles.

18. Students can best develop an understanding of the basic principles of engineering mechanics, not through equations but through the use of experiments and demonstrations.

19. Humans, being logical animals, readily accept, and use instinctively, mathematics which support physical phenomena which we can intuitively predict. Once students understand the fundamental principles of mechanics they will readily accept and use the supporting mathematics.

20. Engineering mechanics 120 should concentrate on the fundamentals of engineering; students can develop an understanding of complex concepts and problems in follow-on engineering courses.

21. It is better to understand simple concepts as opposed to memorizing procedures to complete "plug-and-chug" solutions to more complex problems.

22. Cadets enter engineering mechanics 120 with varied backgrounds and different engineering skill levels.

23. A simple semester-long design project can help illustrate such fundamental concepts as how forces are determined, what results when these forces are applied to a body, force deformations, and material selection.

24. Engineering mechanics 120 should become more of a seminar-type course rather than the lecture and board work style of teaching currently used.

25. Engineering mechanics 120 should be an engineering mechanics application course that focuses on using realistic Air Force examples to illustrate fundamental principles.

26. Engineering mechanics 120 should focus on the engineering problem solution methodology: state the problem and what information is desired, identify important factors and variables, determine the appropriate method to solve the problem, and solve the problem.

27. Introduce students to the engineering vocabulary.[6]

The above observations represent honest opinions of DFEM instructors with varying experience levels in teaching engineering mechanics 120. One common theme echoed in these instructors' views is that this course needs revision and that a clean slate approach to developing an enhanced version of the course offers the best track to producing the highest quality course which addresses their opinions and concerns.

Consequently, after DFEM examined the results of the engineering mechanics 120 time surveys and the department members' opinions, the department formed a group to develop an enhanced engineering mechanics 120 course based on the following objectives:

- to begin with a "clean sheet" without the constraint of keeping the present course content;
- to introduce the engineering method, which can be viewed as a problem-solving philosophy;
- to build on previous or concurrent math, physics, and other science courses;
- to exercise each student's communication skills;
- to examine the introductory engineering courses in other colleges;
- to develop goals for the course and methods to measure success;
- to establish a balance between the process of how engineers think and the content of the course; and
- to be creative, innovative, and forward thinking.[7]

Based on these directives, DFEM plans to develop an engineering mechanics 120 course that focuses on process rather than on content. The department's ultimate goal for an enhanced version of this course is to provide students with an integrated learning experience rather than to learn a set of facts simply to pass an exam. In his proposal to enhance engineering mechanics 120, Colonel Fisher states that

> By making the student central and active in this learning process and by providing a context for the application of the learning to real-life situations, we believe that learner motivation will improve, technical retention will increase, and tomorrow's Academy graduate will be better equipped to thrive with a changing, technical profession.[8]

Enhancing engineering mechanics 120 provides the Air Force Academy with an opportunity to revolutionize engineering education. It offers a tremendous potential for applications in other academic disciplines and in other academic endeavors. Hence, it is imperative that educators develop this enhanced course in a systematic manner and devote adequate and appropriate resources to the course development process to ensure the best possible product. They can apply the assessment guidelines developed in chapter 5 and the principles of assessments outlined in chapter 6 to this development process. Additionally, they can use the information contained in those two chapters, as well as the supporting documentation in the first four chapters of this report, as a way for DFEM to address the many aspects of developing an enhanced version of engineering mechanics 120 and the assessments associated with this new course.

Before I can demonstrate how educators may apply the guidelines and principles outlined in chapters 5 and 6 to the development of this course, I must first state a fundamental assumption with respect to developing a new course such as the enhanced version of engineering mechanics 120. Specifically, if educators wish to develop a new course or enhance an existing one, they undoubtedly will want to assess the effectiveness of the course. Educators will want to determine if the course met its goals, if it is an enhancement of an existing course, and if it is better than the current course. Therefore, educators must develop methodologies to assess student performance not only to determine if the students achieve the desired course objectives but also to enable the educators to evaluate the effectiveness of the course. Assuming this to be the situation for DFEM, then the information contained in this report applies directly to the process of enhancing engineering mechanics 120. What follows, then, examines this course and assessment development process. I raise issues and concerns which I believe DFEM and DFE must answer if they want to develop an effective course that achieves their directives.

Assembling a Cadre

According to the first principle outlined in chapter 6, educators must integrate assessments into the course development process. The first step of this process assembles a cadre to develop the course; as previously mentioned, DFEM has assembled a review committee to serve as the cadre. The department selected members of this committee to represent a cross section of experience and background within DFEM. With Lt Col Bob Pieri as chair, the committee consisted of the following personnel:

1. Col Cary Fisher
2. Col (sel) John Blind
3. Prof Bill Lyons (available in June 1993)
4. Prof Wayne Stinchcomb (departed in June 1993)
5. Maj Bob Hastie
6. Maj Jeff Kouri
7. Maj Steve Whitehouse (available in June 1993)
8. Capt Leslie Blackham
9. Capt Paul Waters (departed in June 1993)
10. Charles Meadows[9]

In addition, the review committee has met with DFE; DFE participants in committee meetings have included Col Joseph Burke, Dr Mary Marlino, and Maj Raymond Caplinger. Lt Col Bob Pieri and several members of DFE have met with representatives of the Training Product Office at Wright-Patterson Air Force Base to develop a statement of work for the initiative to enhance engineering mechanics 120. The Training Product Office has used this

statement to generate a contract to hire a civilian organization to help DFEM develop an enhanced version of engineering mechanics 120.

While DFEM put much thought into selecting the members of the review committee, I wish to address some issues concerning its composition. For example, the decision makers must be members of the committee. These individuals hold the ultimate responsibility for the course and the evaluations of the students in the course, as well as the evaluation of the course itself. As such, their personal involvement plays an essential role in the successful development and conduct of engineering mechanics 120. I assume that the principle decision makers include Colonels Burke, Fisher, Blind, and Pieri. These individuals must be actively involved in developing engineering mechanics 120; lack of their involvement diminishes the importance of this development process and, consequently, the course itself. In fact, an anonymous member of the review committee recently raised the following concern: "We obtained opinions on what is the correct path for the [enhanced version of engineering mechanics 120] to take but we need a decision by someone on which way to go."[10] The decision makers need to play a major role in this course development process.

Two other groups who should be represented in the review committee include the instructors, who will teach engineering mechanics 120, and the evaluators of the course. I assume that some of the members from DFEM, who are on the committee, will teach the enhanced version of engineering mechanics 120 when it is first offered. Since one aspect of assessing the effectiveness of engineering mechanics 120 evaluates student performance, the course instructors must be involved in evaluating the students. However, because the enhanced version of engineering mechanics 120 will differ significantly from the traditional course, the assessment methodologies may in fact be quite foreign to the instructors in DFEM. This departure requires that the instructors receive training to evaluate the information from those assessment methodologies.

In addition to the instructors who will assess the students' performances, the membership of the review committee must include those individuals who will evaluate the effectiveness of engineering mechanics 120. These individuals also must be trained evaluators. However, rather than examine only the students' performances, they also must assess such aspects of the course as the teaching style, learning environment, and materials the instructors use to determine if the course is achieving the specified goals and objectives.

A final issue concerning the composition of the review committee deals with the individuals who will use the information generated from the assessments of the enhanced version of engineering mechanics 120. At first glance, I assume that the individuals in this group include the instructors who evaluate the students' performances, as well as the educators who will assess the effectiveness of engineering mechanics 120. These educators will include not only the trained evaluators but also the decision makers, since the latter group must use the results of the assessments to determine the fate of the course. Another group of educators who will use the assessments to examine

the effectiveness of the course will be those individuals who plan to conduct studies of the enhanced version of engineering mechanics 120. Finally, the instructors in the follow-on courses which require engineering mechanics 120 as a prerequisite comprise an important group of information users.

Based on the current committee membership, the decision makers are well represented on the committee. Dr Marlino and several of the members of DFEM who serve on the committee are well qualified to conduct controlled and thorough studies of the effectiveness of the enhanced version of engineering mechanics 120. Unfortunately, I contend that the information users have inadequate representation on the review committee. Specifically, the instructors of the courses which require engineering mechanics 120 as a prerequisite have negligible involvement on the review committee.

Engineering mechanics 120 is a prerequisite for the core engineering courses aeronautical engineering 215, astronautical engineering 320, and engineering 410. More importantly, according to Colonel Fisher, "as the *first* engineering core course, we can set the stage for the other follow-on engineering courses."[11] In fact, a number of advanced engineering courses require engineering mechanics 120 as a prerequisite or a corequisite. These courses include civil engineering 264, civil engineering 330, civil engineering 350, civil engineering 361, civil engineering 362, engineering mechanics 290, engineering mechanics 320, and engineering mechanics 331. Further, biology 420, a biomechanics course, requires engineering mechanics 120 as a prerequisite. Therefore, the instructors of these courses and the instructors of the other core engineering courses must have representation on the review committee.

The logic behind this assertion has a straightforward meaning. Instructors who build on the information presented in engineering mechanics 120 must have a say in what specific objectives the engineering mechanics 120 instructors will teach. Otherwise, there is no reason that this course should serve as a prerequisite to follow-on courses, and DFEM should not link engineering mechanics 120 to other engineering courses. However, if the committee were to include every instructor with an interest in engineering mechanics 120, then the committee would become too large and unmanageable. The review committee solved this dilemma by drawing up plans to survey those instructors to determine what knowledge, skills, and higher level cognitive abilities they expect students in engineering mechanics 120 to master. The review committee will then use this information to develop the course goals and objectives.

Establishing the Goals for Engineering Mechanics 120

Once educators have assembled the review committee, the committee should develop the goals for the enhanced version of engineering mechanics 120. This is a crucial undertaking, as the goals for the course establish the direction the committee will pursue in developing engineering mechanics 120. The current

committee already has devoted much time and energy into establishing these goals. They have outlined a statement of their goals for this course:

> The first engineering course is designed to provide beginning students with a fundamental understanding of and an appreciation for engineering and engineering methods. The course emphasizes the systematic approach to basic problem solving, the understanding and description of physical situations, and communications. The course utilizes and complements other concurrent courses (math, science, and English) and also provides a firm foundation for learning in technical courses which follow. The fundamentals of engineering will be developed through the application of statics, dynamics, and mechanics of materials.[12]

The above quote illustrates the ambitious goals the review committee has established for the enhanced version of engineering mechanics 120. The committee can summarize these goals concisely by stating that, by the conclusion of engineering mechanics 120, students will:

1. fundamentally understand engineering;
2. fundamentally understand engineering methods;
3. learn the systematic approach to basic problem solving;
4. understand and describe physical situations;
5. exercise communication skills;
6. use knowledge learned in previous and concurrent math, science, and English courses;
7. develop a firm foundation for learning in follow-on technical courses;
8. understand the fundamentals of statics;
9. understand the fundamentals of dynamics; and
10. understand the fundamentals of mechanics of materials.[13]

Of course, the review committee derived these goals from the directives specified in the proposal to enhance engineering mechanics 120. While the goals comprise an impressive list, it omits several concerns that the committee must address to ensure that these goals are clear, realistic, achievable, and observable.

An adage comes to mind when I ponder the task of assessing the effectiveness of engineering mechanics 120 to determine if it will achieve the above goals; that is, if it isn't broken, then don't fix it. Other than the fact that DFEM wishes to enhance it, what is wrong with the current version of engineering mechanics 120? Is a new course necessary, and, if so, why? Is engineering mechanics 120 now achieving its goals? In fact, what are the current goals for this course, and how do they differ from the goals stated above? How has DFEM previously assessed the effectiveness of engineering mechanics 120? Is it possible that the course itself is effective but that the assessment methodologies are not appropriate or that educators have not properly employed them so that evaluators can adequately determine the success of the course?

Now, consider these questions in another light. After examining the goals the review committee has established, one major flaw surfaces: educators want engineering mechanics 120 to provide everything to everybody. Given

the simple constraint that this is a single semester-long course, it cannot accomplish every person's desires. Therefore, the goals the review committee has established violate the first principle of assessment; that is, the committee must establish clear, realistic, achievable, and observable goals. The committee members cannot design engineering mechanics 120 to achieve their goals and at the same time address every instructor's desires and concerns. Because most instructors in DFEM are engineers first and educators second and are therefore inexperienced in developing undergraduate courses, the committee must solicit assistance from the professional educators in DFE to develop the appropriate goals for this course. Examining several of the existing goals and their impact on the stated purpose of this course serves to illustrate the importance of establishing clear, realistic, achievable, and observable goals for engineering mechanics 120.

For example, the current version of engineering mechanics 120 provides students with a basic understanding of statics and strengths of materials. Several instructors in DFEM believe that certain topics they introduce in the current version of engineering mechanics 120 confuse the students and should be offered in more advanced engineering courses; three-dimensional equilibrium is one such topic.[14] Eliminating those topics the instructors determine are not essential for the students to understand engineering would free up blocks of lessons so the instructors could cover other topics like dynamics and engineering design. Granted, eliminating, for example, three-dimensional equilibrium would reduce the confusion each student has in understanding this difficult concept; unfortunately, it would only open up four lessons to introduce dynamics and engineering design. In addition, astronautical engineering 320, a core course, requires students to understand three-dimensional vectors. Students would be at a disadvantage if instructors in engineering mechanics 120 did not expose them to this concept, unless instructors in a physics course or an astronautical engineering 320 course could introduce it.[15] In that case, would astronautical engineering 320 no longer require engineering mechanics 120 as a prerequisite? If so, how would that fit into the concept of using engineering mechanics 120 to set the stage for the follow-on engineering courses at the Air Force Academy? The review committee must therefore determine the impact of eliminating topics covered in the current version of the course.

If, conversely, the committee eliminated topics like three-dimensional equilibrium from the course content, a popular sentiment among DFEM faculty members is that they should use the open lessons to introduce students to the engineering design process.[16] Engineering design is, according to the American Board of Educational Technology criteria,

> The process of devising a system, component, or process to meet desired needs. . . . The engineering design component of a curriculum must include most of the following features: development of student creativity, use of open-ended problems, development and use of modern design theory and methodology, formulation of design problem statements and specifications, consideration of alternative solutions, feasibility considerations, production processes, concurrent engineering design, and detailed

> system descriptions. Further, it is essential to include a variety of realistic constraints, such as economic factors, safety, reliability, aesthetics, ethics, and social impact.[17]

As the above definition implies, engineering design is a complex topic. As such, instructors cannot adequately address it in a few lessons. In fact, instructors in engineering mechanics 290 devote the entire semester to introducing engineering mechanics and engineering sciences students to the engineering design process. While it concentrates on critical thinking and problem solving, engineering mechanics 290 cannot adequately introduce students to the many facets of engineering design. Devoting several lessons in engineering mechanics 120 to engineering design would only serve to survey the discipline and may, in fact, only confuse the students. Therefore, the review committee must decide why they want to integrate engineering design into engineering mechanics 120 and what aspects of design they want to concentrate on if this topic is still to be an essential part of the enhanced version of the course.

Similarly, engineering mechanics 120 may not be the proper forum for introducing students to dynamics. Members of the review committee have suggested devoting one lesson in the enhanced version of engineering mechanics 120 to addressing Newton's Second Law by introducing acceleration into the equilibrium equation instructors teach the students.[18] Unfortunately, a pitfall to this approach is that instructors would only give "face time" to dynamics without providing students with a meaningful introduction that allows them to "appreciate" the complexity of this subject. Engineering mechanics 320 is the required dynamics course for astronautical engineering, aeronautical engineering, engineering mechanics, and engineering sciences majors. This semester-long course exposes students to dynamics, which, like engineering design, is a complex discipline. Devoting part of engineering mechanics 120 to dynamics would serve only to confuse the students, much like the topic of three-dimensional equilibrium already does. Besides, the physics department presents students with the fundamentals of dynamics in their core courses.

I have discussed the problems associated with using engineering mechanics 120 to introduce students to engineering design and dynamics to demonstrate that the goals and desires for the enhanced version of engineering mechanics 120 expressed by DFEM may be ambitious. Again, I emphasize that the review committee must establish clear, realistic, achievable, and observable goals for this course. For example, the review committee views using the enhanced version of this course to develop within each student an intuitive understanding of the fundamental principles of engineering mechanics. While this goal may be clear, realistic, and achievable, it is not observable. What assessments can the committee develop to determine if students achieve this goal?

It may appear that I have devoted too much time to the principle of establishing clear, realistic, achievable, and observable goals. However, these goals are critical to developing a worthwhile and effective course, as well as the associated assessments. In developing the course goals, the review committee must critically examine and incorporate the opinions and viewpoints of the

instructors in DFEM and of the instructors of those courses which require engineering mechanics 120 as a pre- or corequisite. The resulting goals will undoubtedly compromise the desires of individual review committee members for the good of the whole committee. The committee has recognized the need to limit the goals to those which the students can achieve in one semester. The members of the committee also have recognized the requirement to limit the fundamental engineering principles they wish the students to study. It would be an injustice to cover statics, dynamics, materials, and engineering design in engineering mechanics 120 if the committee wants the students to develop a fundamental understanding of the disciplines. However, committee members must first define what they mean by a "fundamental understanding" of a particular discipline.[19]

I do not propose to develop course goals for the enhanced version of engineering mechanics 120, but I wish to offer suggestions which may clarify what I mean by establishing goals which are clear, realistic, achievable, and observable. These suggestions are, in fact, derived from the statements of the review committee members which specify their views of what engineering fundamentals the instructors should address in an enhanced version of engineering mechanics 120.[20] A more constrained set of goals for engineering mechanics 120 could include demonstrating an understanding of:

1. Newton's laws of mechanics by describing the physical behaviors of selected systems of interest to the Air Force;
2. the fundamental concepts of statics by applying two-dimensional equilibrium to engineering problems;
3. the basic properties of selected engineering materials by applying these properties to engineering problems; and
4. selected terms related to engineering by defining and using them appropriately.

These goals have characteristics which satisfy the first principle of assessment. They are clear. Students who take engineering mechanics 120 can achieve them. At the same time, the stated expectations of the course are realistic; the goals do not seek to offer everything to everybody. Note that the goals make no mention of addressing dynamics and engineering design, but they still retain the current engineering mechanics 120 focus of statics and strengths of materials. While the review committee may choose to focus on different engineering fundamentals, members must keep the focus narrow enough for the students to understand them. Finally, the goals are observable. From these goals the review committee can develop objectives which are measurable and, therefore, assessable.

Developing the Objectives for Engineering Mechanics 120

Once the members of the review committee establish the goals for the enhanced version of engineering mechanics 120, they can develop the associated

objectives. Like the goals, these objectives must be clear, realistic, and specific. They must also be measurable, since evaluators will assess student performance to determine if the students achieve the objectives. Therefore, the objectives must define the required performances of the students. Terms describing measurable objectives include perform, define, use, complete, describe, and delineate. Such terms as analyze, recognize, comprehend, understand, and know are not measurable. Again, because most DFEM instructors are engineers first and educators second, the committee members must rely on the expertise of professional educators in DFE to help them establish the appropriate objectives for engineering mechanics 120.

For example, while instructors may want students to comprehend a specific topic, they cannot measure a level of comprehension. This statement does not imply that student comprehension should not be a course objective; indeed, engineering mechanics 120 must help develop a student's higher order cognitive skills and abilities. However, the review committee members must explicitly state a comprehension objective in measurable terms; students can demonstrate comprehension of a topic by applying it to a new situation or describing physical phenomena related to that topic. In other words, the objectives must contain specific criteria which evaluators can assess, and they must contain constraints or specific circumstances during which evaluators can conduct the assessments. Finally, the review committee must identify the level of learning for each objective and specify whether it is short-term performance in class or long-term transfer of knowledge to related courses.

Since the objectives and levels of learning depend on the goals for engineering mechanics 120 and since I assert that the current goals the review committee has established need refining, I cannot comment on the current objectives the committee has specified. Again, I do not wish to appear presumptuous enough to develop the objectives for the review committee, but I can offer suggestions of the types of objectives which satisfy the criteria of being clear, realistic, specific, and measurable. I will therefore focus on one of the goals I specified in the previous section of this chapter and establish a set of associated objectives. The specific goal I will examine is the fourth one: to demonstrate an understanding of the basic properties of selected materials by applying these properties to engineering problems.

Naturally, the objectives I develop must relate directly to this goal. The review committee already has established several such objectives which I can express in clear, concise, and measurable terminology:

1. identify, differentiate, and solve for axial, bending, shear, and torsional stresses;
2. describe specific material properties and how they effect aircraft; and
3. describe the physical response of a material under axial loading.[21]

Note, from the above example, that each objective uses an active verb which communicates that an evaluator can assess the objective. Each objective also contains specific and measurable criteria. The four objectives illustrate a mix of behavioral and cognitive learning objectives. Specifically, the first objective

is behavioral because it requires knowledge of stresses; the review committee can develop norm-referenced assessments to test student achievement of this objective. The remaining three objectives are both behavioral and cognitive. They are behavioral because they require student knowledge of material properties and the behavior of materials. The objectives are also cognitive since they require the students to understand different aspects of material properties or behaviors and then demonstrate their understanding through specific applications. Even though the last four objectives are more involved than the first, the review committee can develop, for example, norm-referenced, criterion-referenced, and performance assessments to determine student achievement of the objectives.

Once they develop the specific learning objectives for the course, the members of the review committee can establish desired levels of learning for the objectives. These learning levels relate to the knowledge, skills, and higher order cognitions the committee wants students to grasp for each objective. Consequently, the committee must tie these learning levels to the goals that specify what the result of the course should be. Since engineering mechanics 120 seeks to introduce students to the fundamentals of engineering, I assume that the levels of learning associated with the above objectives should be comprehension and application learning levels related to the long-term transfer of knowledge to follow-on courses. Thus, the instructors of those courses that require engineering mechanics 120 as either a pre- or corequisite must specify the knowledge, skills, and higher order cognitions they require the students to retain. However, if the follow-on courses are not part of the core curriculum, then why should educators require students in the nontechnical tracks to achieve these objectives? Further, if instructors do not require the learning objectives for follow-on courses, then why should students need to achieve them? If, instead, the objectives satisfy the review committee's goal of having the student fundamentally understand engineering and engineering methods that follow-on courses do not require, are the objectives and the associated levels of learning appropriate or even needed? In other words, why, for example, require that a student with a nontechnical major identify, differentiate, and solve for axial, bending, shear, and torsional stresses? The review committee must therefore establish objectives and levels of learning that are not only clear, specific, and measurable but also are realistic and appropriate for the students.

Comparing the Existing and Enhanced Versions of Engineering Mechanics 120

Once the members of the review committee establish the behavioral and cognitive objectives and the associated levels of learning for the enhanced version of engineering mechanics 120, they can develop the course and the appropriate methodologies to assess student performance and the effectiveness of the course. While I will continue to apply the guidelines and

principles I discussed in the previous chapters to engineering mechanics 120, I must first emphasize a significant disparity which exists between the thrust of my report and this course.

Specifically, my report establishes guidelines and principles for developing methodologies to assess student performance in different learning environments so that evaluators can compare the effectiveness of different teaching styles. Unfortunately, the review committee cannot compare in its entirety the existing engineering mechanics 120 to its enhanced version. One principle stated in chapter 6 has already been violated; that is, the committee has not established identical goals, behavioral and cognitive objectives, and associated levels of learning for the two courses. Therefore, the review committee cannot discern, based on any qualitative and quantitative studies, which version of the course is more effective; the two courses are fundamentally different. However, the review committee can compare isolated segments of the two courses for those goals and objectives which remain the same.

In addition, the committee can generate qualified conclusions relating to students' opinions of the enhanced version of engineering mechanics 120 when compared to other courses the students are enrolled in. The committee also can make general conclusions based on instructors' opinions of the course when compared to their impressions of the existing engineering mechanics 120 course or other courses which they teach. Conversely, the committee cannot compare student test results from both courses, nor can they quantify such a statement as students spend 50 percent more time studying for the enhanced version of the course than they do for the existing course. Such comparisons are impossible, since the two courses have different goals, objectives, and levels of learning.

Developing Engineering Mechanics 120 and Its Assessment Methodologies

Assuming the review committee has established comprehensive behavioral and cognitive objectives, it can develop lessons to address each objective. Since the objectives relate to the goals for engineering mechanics 120, the review committee must establish methodologies to assess if the students in the course actually achieve these goals and objectives; students who successfully achieve the objectives also have attained the goals of the course. The success or failure of the enhanced version of the course should depend on how well the students achieve the course objectives; hence, the assessment methodologies play a critical role in determining the success of this course.

The members of the review committee must be imaginative when they develop the methodologies to assess student achievement of the course objectives. For example, some objectives may relate to the students' motivational levels or attitudes; evaluators can measure these objectives by using surveys and questionnaires. Other objectives may require extensive and varied assessments to determine the level of the students' performances. The

committee must therefore develop realistic, thorough, diverse, and accurate methodologies to generate a comprehensive assessment of student performance, and, in the end, a comprehensive assessment of the effectiveness of the enhanced version of engineering mechanics 120.

Hence, the review committee must integrate assessments into the course's lesson plans. The committee members, when developing the lessons, must determine how the information they want to present relates to the course objectives and what assessments they can use to determine if the students learned and understood the material. Since the assessments the review committee will develop depend on the lessons, material presented, and course objectives, I cannot suggest specific assessments to use. However, I can enlighten the committee on facets of different assessments as they develop their assessment methodologies.

Assessing Student Performance

I refer to assessment methodologies in the plural form, because, as one of the principles in chapter 6 states, the review committee must develop a blend of assessments to provide a complete picture of a student's performance in the enhanced version of engineering mechanics 120. Learning is a dynamic process, and this course should help students to develop and expand several levels of their learning. Knowledge, comprehension, and application are three such levels.

One assessment methodology, the portfolio, provides a means to develop this comprehensive picture of a student's performance, as it can be a balance between behavioral and cognitive assessments. Behavioral assessments include traditional methods that evaluators use to determine a student's mastery of knowledge; such quantitative assessments include multiple choice examinations and standardized tests.[22] Chapter 4 detailed advantages and disadvantages of these types of assessments. I believe that the review committee should employ them to determine if students do indeed grasp the behavioral objectives of engineering mechanics 120; the committee can also use these assessments to determine how much knowledge students retain after they leave the course. The committee must, however, be imaginative in developing these behavioral assessments to ensure that they do not test a student's test-taking ability and that they test relevant information only.

Evaluators can apply the same caution to developing cognitive assessments; in fact, this caution may be more critical since evaluators usually find cognitive assessments more difficult to develop and administer than behavioral assessments. Cognitive assessments are more qualitative in nature, measuring the higher order learning skills and abilities of a student.[23] Again, chapter 4 discussed such cognitive assessments as authentic and performance assessments. Quite often, educators can use the written assignments from cognitive assessments to determine a student's comprehension of information effectively, as the student must retrieve the information, relate it to different concepts, and then externalize it so that other people can understand what the student wants to communicate.[24]

Educators cannot accomplish such cognitive assessments, as well as behavioral assessments, only once or twice during the semester. They must repeat the assessments periodically to see how students develop their behavioral and cognitive knowledge, skills, abilities, and comprehension as the semester progresses. Repeated cognitive assessments, for example, can show how a student's critical thinking develops with respect to engineering as instructors expose him or her to the various engineering fundamentals during engineering mechanics 120. Assessments do not necessarily need to be formal; such short and impromptu assessments as one-minute papers can help the instructor gauge individual and class progress towards the desired course objectives. I must emphasize, though, that each behavioral and cognitive assessment must be directed towards measuring a specific objective for the enhanced version of engineering mechanics 120.

An important aspect of this measurement process is establishing a baseline performance standard. While periodic assessments determine how a student progresses towards achieving the objectives, the review committee must establish a baseline measure to help determine each student's overall performance and to provide a means to measure his or her cognitive growth during the semester. Chapters 1 and 5 discussed how educators can develop such baseline measures. Students enter engineering mechanics 120 with different abilities and knowledge levels; some students may not exert the same effort as others to achieve the course objectives. The effectiveness of a particular teaching style depends on the level of performance each student achieves with respect to the learning objectives of engineering mechanics 120. The review committee can use a baseline measure to help instructors encourage those students with advanced abilities to pursue additional tasks and assignments so that they can get the most out of the course. The baseline measure also provides a performance standard which allows instructors to determine how far each student has progressed. This standard helps the instructor assign the appropriate grades at the end of the course.

Assessing the Effectiveness of the Course

As noted above, a baseline measure helps instructors to determine each student's progress towards achieving the objectives of the enhanced version of engineering mechanics 120 so that they can assign grades after the students have completed the course. The review committee must therefore design the assessments with the understanding that the prime motivator to assess student performance is to assign grades after the students complete this course.

Student performance is just one way to determine the effectiveness of the enhanced version of engineering mechanics 120; evaluators also need to assess other factors if the review committee requires a thorough scrutiny of the course's effectiveness. Certainly one factor is the long-term impact of the course on a student's retention of the information presented. I must assume that the review committee would want to determine what aspects, knowledge, skills, and higher order cognitions related to engineering mechanics 120 each student carries to

other courses. The committee members should establish what threshold of basic knowledge, skills, abilities, and comprehension they want the students to retain regardless of the future courses they take. Consequently, the committee must track the students who take the enhanced version of this course and periodically assess them after they have completed the course. The assessments can resemble those used during the course, but they should assess the knowledge, skills, and higher order cognitions the committee determines are essential, either to future courses or to the students' long-term academic growth.

The review committee can also assess the effectiveness of the enhanced version of engineering mechanics 120 by using student surveys and questionnaires. Chapters 4 and 5 discussed in detail various ways to develop and administer the surveys and questionnaires. The committee members must establish the appropriate student characteristics to help them determine how the course influences each student's attitudes and motivation levels. Surveys and questionnaires offer appropriate tools to handle this task. As I stated in one principle in chapter 6, the review committee must administer these evaluation forms at the beginning and the end of engineering mechanics 120, as well as periodically during the semester. If the committee members administer the forms to students before the semester begins, they can use the generated information to select the students to enroll in the enhanced version of engineering mechanics 120, especially if the committee wants to conduct studies to examine the effectiveness of the course.

A final factor in assessing the effectiveness of the enhanced version of engineering mechanics 120 is each instructor's opinions and attitudes towards the course. Two principles in chapter 6 outlined how the review committee can assess the instructors' impressions of this course. The first task required the instructors themselves to objectively assess the effectiveness of specific lessons with respect to how the lessons and the associated behavioral and cognitive objectives impacted student performances, attitudes, and motivation levels.

The second principle recommended having the instructors' colleagues assess the effectiveness of engineering mechanics 120. The colleagues, who must be trained as evaluators, should periodically attend the classes and observe the interactions between the students and the instructors. They should observe how such aspects of the lesson as the teaching style and the environment influence the students' performances, attitudes, and motivation levels. The colleagues would not assess the instructors themselves but would provide objective assessments of the effectiveness of the course.

The various assessment methodologies I just described provide the review committee with the appropriate information to assess the effectiveness of the enhanced version of engineering mechanics 120. The committee must, however, realize that accurate and realistic assessments of teaching styles and learning environments require long-term commitments from each member of the committee. The committee simply cannot spend one semester changing the course and another administering the new course and still expect immediate feedback on the effectiveness of the course Granted,

various aspects of the enhanced version of engineering mechanics 120 will permit immediate feedback, and the committee can alter or fine-tune those aspects as required. But instructors must teach this course for several semesters before the review committee can assess the effectiveness of the course realistically.

Summary

In this chapter I have applied the guidelines and principles summarized in chapters 5 and 6 to a course and integrated assessment development process. In addition, I have raised issues and concerns with respect to developing an enhanced version of engineering mechanics 120. I did not seek to develop the course and the assessments themselves but to highlight various aspects of the course development process and discuss the need to integrate assessments in this process. I emphasized the requirement to establish clear, realistic, achievable, and measurable course goals and the associated behavioral and cognitive goals for two reasons. First, and most important, the effectiveness of engineering mechanics 120 depends on formulating and delineating the goals and objectives of the course; the review committee must know specifically what they want out of the course before they can develop it. Second, effective and appropriate assessments depend on the goals and objectives. How can you assess something if you don't know what it is? In conjunction with establishing course goals and objectives, educators who wish to develop or enhance an educational program must develop a blend of methodologies which will allow them to thoroughly and accurately assess student performance and, therefore, the effectiveness of the course.

This chapter has reflected some personal opinions and is based on my experience as a student and an instructor. I offer this chapter with a sincere desire to help the Department of Engineering Mechanics to develop an introductory engineering course that will benefit cadets.

Notes

1. Col Cary A. Fisher, professor and head, Department of Engineering Mechanics, to Col Joseph G. Burke, Directorate of Education, letter, subject: Proposal to Participate in a Joint DFE/Department Course Development Initiative, with attachment, 31 December 1992.
2. Ibid.
3. Ibid.
4. Ibid.
5. Ibid.
6. Anonymous opinions of Department of Engineering Mechanics instructors on the enhanced engineering mechanics 120 development process.
7. Fisher letter.
8. Ibid.
9. Ibid., 3.

10. Anonymous opinions of Department of Engineering Mechanics instructors on the enhanced engineering mechanics 120 development process.

11. Fisher letter.

12. Fisher letter attachment, 2.

13. Ibid.

14. Anonymous opinions of Department of Engineering Mechanics instructors on the enhanced engineering mechanics 120 development process.

15. Ibid.

16. Ibid.

17. Fisher letter.

18. DFEM review committee meeting notes, 14 October 1992.

19. Lt Col Bob Pieri, chairman of the engineering mechanics 120 review committee, Department of Engineering Mechanics, US Air Force Academy (HQ USAFA/DFEM), personal interview with author, 1 June 1993.

20. Ibid.

21. DFEM review committee meeting notes, 7 October 1992.

22. Dr Kathleen Hannafin, visiting professor, Department of Behavioral Sciences and Leadership, US Air Force Academy (HQ USAFA/DFBL), interview with author, 19 January 1993.

23. Ibid.

24. Ibid.

Chapter 8

Conclusions

This study has developed guidelines and principles for educators when they develop methodologies to assess different styles of teaching. In this report I reviewed various traditional teaching methods and innovative educational technologies. In addition, I summarized studies that assess the effectiveness of teaching in the traditional manner and teaching using educational technologies as teaching aids. I also examined diverse methodologies which assess teaching effectiveness. Based on this supporting information I outlined specific guidelines and principles which educators must consider to develop or enhance their undergraduate courses and the associated assessment methodologies. Finally, I examined a case study which implemented these basic guidelines and principles.

The information contained in this study is not limited to undergraduate institutions. Rather, it applies to educational institutions, ranging from elementary and secondary schools to training, graduate, and professional education institutions. In fact, anyone who teaches a single student or a group of students should examine the guidelines and principles I have developed here. These guidelines and principles provide a fundamental direction for educators to follow as they build courses that not only satisfy their goals but also provide effective learning environments for their students.

As I repeatedly emphasized in this study, before educators can develop courses and assessment methodologies, they need a clear and focused idea of what they want as the end state of the course. What knowledge, skills, abilities, understanding, comprehension, and attitudes do they want their students to gain from the course? What performance measures will they want to use to determine if the students have achieved the educators' goals? How will they measure the students' performances? How will they determine the effectiveness of the course? For that matter, what do they mean by the "effectiveness" of the course?

Educators can find these basic questions quite difficult to answer, but they must develop some answers if they desire direction and purpose for their courses. The assessment methodologies which educators develop for the course must have two objectives: to assess student performance and to assess the effectiveness of the teaching style. The assessments must be integrated into the course, and they must relate to the course goals. Consequently, educators must ensure that the goals are achievable and observable so that they can develop clear, specific, realistic, and measurable objectives. The assessments then will measure student performance and teaching style

effectiveness according to how well the students achieve those objectives. This relationship between the goals and assessment methodologies underscores the importance of concentrating on course development.

Educators generally do not emphasize the establishment of course goals and objectives before they begin to develop a course. Consequently, many courses fail to accomplish what educators intended. Similarly, educators do not incorporate the development of a comprehensive assessment methodology as they build the course and, as a result, they have an incomplete picture of each student's achievements in the course and the effectiveness of the course. Most assessments are piecemeal and uncoordinated efforts to measure student performance in isolated segments in the course.

While many innovative assessments exist, educators either improperly employ them or fail to integrate them in a comprehensive package of assessments. Such assessments as performance and authentic assessments do indeed challenge a student's higher order learning skills, abilities, and understanding, but educators must not exclusively use them to assess student performance. The more traditional norm- and criterion-referenced assessments still yield invaluable information for educators; in fact, they can assess quantitative aspects of student performance, something qualitative assessments do not measure. Hence, a comprehensive assessment methodology must include both qualitative and quantitative assessments.

Educators must employ these assessments periodically during the course; assessing a student's performance only once or twice fails to represent his or her performance and achievement accurately. Too often, educators base student performance on quantitative measures and do not examine such achievement parameters as motivation, enthusiasm, comprehension, and long-term retention of knowledge, skills, abilities, and understanding. From an educator's viewpoint a student performs well in a course when he or she excels on examinations. However, that student might be a good test-taker and have a poor attitude in class; he or she may study only to pass the examinations and fail to retain anything from the course. Educators must ensure that their assessments can measure the aspects of student performance that they deem appropriate for the course; they must base these measures on the course's goals and objectives.

Assessments are essential tools to determine the effectiveness of a particular teaching style; however, they comprise only part of the process of assessing teaching effectiveness. Comprehensive research studies must incorporate the results of these assessments so researchers can make confident and unbiased conclusions concerning the teaching styles. While educational researchers have conducted many studies that examine the effectiveness of teaching styles, this field of research is ripe for further exploration. Future studies should concentrate on the long-term effects of different teaching styles on students. The true worth of a particular style of teaching lies in the quality and quantity of knowledge, skills, abilities, comprehension, and experiences the student retains from the course he or she has completed. Educators should conduct research studies over a long period

of time and incorporate large sample sizes of students and assessments of the students' retention and comprehension of material. Researchers must examine a number of courses in different academic disciplines before they can generalize the results of their studies of teaching effectiveness; a particular teaching style that is effective for one course may be ineffective for another course. Finally, educators must commit themselves to a long-term research effort if they expect to generate viable studies.

Although my research has focused on assessments, I have addressed various styles of teaching; specifically, traditional teaching methods and teaching using innovative educational technologies. Granted, educational technology opens an exciting spectrum of teaching styles and has tremendous potential to enhance the learning process. However, traditional teaching methods are still viable forms of education. Like assessments, some teaching styles are better suited for particular courses or disciplines than others; educators must use those styles which will make their courses the most effective. In many situations, the "best" teaching style is actually a blend of different styles. Variety may enhance a course more than using a single teaching style, but the teaching styles must complement, rather than detract from, the learning process. Educators cannot introduce innovative educational technologies and still rely on a traditional teaching method; these technologies are effective only if instructors use them appropriately. This statement holds for traditional teaching methods as well. A learning environment will be effective if educators adapt one or more styles of teaching which suit the instructor, the students, and the course.

Education is a dynamic and complex process. Educators must constantly improve on the way they teach students. Perhaps innovative educational technologies resemble vehicles to enhance a student's learning process. However, traditional teaching methods still have a place in the educational system. What educators must accomplish, then, is to employ traditional teaching methods and teaching using innovative educational technologies effectively so that the learning process produces a student who has achieved the course goals and has a positive attitude towards the subject matter and the learning process. Effective assessments help the educators to determine if they are successful in their mission. This is a challenge, but the payoff is well worth the effort.